Immigration and Local Business Dynamics: Evidence from U.S. Firms

by

Parag Mahajan
University of Delaware

CES 21-18 August 2021

Abstract

This paper finds that establishment entry and exit—particularly the prevention of establishment exit—drive immigrant absorption and immigrant-induced productivity increases in U.S. local industries. Using a comprehensive collection of confidential survey and administrative data from the Census Bureau, it shows that inflows of immigrant workers lead to more establishment entry and less establishment exit in local industries. These relationships are responsible for nearly all of long-run immigrant-induced job creation, with 78 percent accounted for by exit prevention alone, leaving a minimal role for continuing establishment expansion. Furthermore, exit prevention is not uniform: immigrant inflows increase the probability of exit by establishments from low productivity firms and decrease the probability of exit by establishments from high productivity firms. As a result, the increase in establishment count is concentrated at the top of the productivity distribution. A general equilibrium model proposes a mechanism that ties immigrant workers to high productivity firms and shows how accounting for changes to the firm productivity distribution can yield substantially larger estimates of immigrant-generated economic surplus than canonical models of labor demand.

Keyword*:* Immigration, Business Dynamics, Job Creation, Productivity, Firm Heterogeneity

JEL Classification: J23, J61, L11, F22

[*] I would like to thank John Bound, Dean Yang, Jeffrey Smith, Hoyt Bleakley, and Jagadeesh Sivadasan for invaluable advice and support. I would also like to thank Pia Orrenius, Gianluca Orefice, Joan Llull, Sebastian Ottinger, Kirill Borusyak, Peter Hull, Andrei Levchenko, Nicolas Morales, Steven Haider, Brian Kovak, Emek Basker, Ben Lipsius, Laurien Gilbert, Dhiren Patki, Ariel Binder, Max Risch, and seminar participants at the University of Michigan, U.S. Census Bureau CES, University of Wisconsin, Colby College, University of Delaware, RAND Corporation, Federal Reserve Board, University of Virginia, FDIC, Federal Reserve Bank of Dallas, University of Colorado–Denver, Wesleyan University, U.S. Treasury OTA, Michigan State University, CUNY Baruch College, MEA Annual Conference, SOLE Annual Meetings, APPAMFall Research Conference, SEA Annual Meetings, and NBER Immigrants and the U.S. Economy conference for helpful comments and suggestions. Finally, I would like to thank Joseph Ballegeer and J. Clint Carter for help and support in accessing data and disclosing results and Jarcy Zee for data visualization assistance. Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2105. (CBDRB-FY21-P2105-R9085)

1 Introduction

The Census Bureau estimates that by 2030, immigration will overtake natural increase as the primary driver of population growth in the United States.[1] This far-reaching demographic change will translate into a workforce that increasingly relies on the foreign-born, magnifying the need for a comprehensive understanding of how they are absorbed into labor markets and ultimately shape industries. Recent advances to data and theory have dramatically expanded our insight into the role of the firm in mediating these processes, with a particular focus on the form and choice of production technique.[2] Nonetheless, most of this literature has either implicitly or explicitly restricted its attention to representative firm models of production that, by definition, do not feature differences across firms in input use or total factor productivity. Moreover, the empirical work that has studied employer-level responses to immigration has largely centered on continuing firms in non-U.S. settings.[3]

In contrast, broader study of the U.S. economy—the largest immigrant destination in the world—finds that business entry and exit dynamics are crucial drivers of job creation and productivity growth, particularly when entry is accompanied by the exit of less productive businesses.[4,5] Furthermore, immigrant workers appear to be especially active in changing business entry and exit decisions: immigrants have a higher propensity to start firms than natives,[6] are relatively more likely to work for new firms,[7] and prevent establishment exit in the short run.[8,9] When viewed through this lens, the question of how local industries generate enough jobs to absorb immigrant inflows naturally leads us to consider business entry and exit. When we open the door to differences in productivity and input use across firms, such entry and exit introduces a new channel through which immigrants alter an industry: by changing its composition of businesses.

Using 40 years of unique, confidential data from the U.S. Census Bureau, this paper presents a comprehensive analysis of how immigrant-induced labor supply shocks impact business entry and exit decisions. I measure immigrant worker inflows into a local industry—defined as a pairing between one of 722 commuting zones and 41 industry groups—using demographic data that includes all survey responses to the 1980, 1990, and 2000 Decennial Census Long Forms and 2005–2019 American Community Surveys. I test how these inflows affect establishment entry and exit dynamics, which are primarily measured using the Longitudinal Business Database, an establishment-level panel dataset that contains administrative records from the near-universe of the U.S. private

[1] Natural increase is defined as births minus deaths of natives. See Vespa and Armstrong (2018).

[2] See, e.g., Lewis (2005, 2012); Clemens et al. (2018); Dustmann and Glitz (2015); Peri (2012); Mitaritonna et al. (2017); Lewis (2011); D'Amuri et al. (2010); Peri and Sparber (2009); Gonzalez and Ortega (2011).

[3] See, e.g., Mitaritonna et al. (2017); Brinatti and Morales (2021).

[4] See, e.g., Foster et al. (2008) and Bartelsman and Doms (2000).

[5] A vast majority of firms in the U.S. are single-unit. However, because several important employers in the U.S. are multi-unit, it is important to distinguish between a firm and an establishment. I use the term "business" to encompass firms and establishments more generally.

[6] See, e.g., Fairlie and Lofstrom (2015); Kerr and Kerr (2018); Azoulay et al. (2020).

[7] See, e.g., Kerr and Kerr (2016).

[8] See Orrenius et al. (2020).

[9] This paper will use the term workers to encompass both the self-employed and employees.

sector from 1976 onward. To resolve endogeneity concerns endemic to the study of immigration on economic outcomes, I develop a new shift-share instrument that incorporates bilateral emigration data from non-U.S. OECD member nations to isolate exogenous migration pushes from over 120 sending countries. These exogenous pushes are distributed into specific local industries using prior compatriot locational choices and contemporaneous compatriot industry choices in other Census Regions.

This setup allows me to develop several novel results that demonstrate the central mediating role of establishment entry and exit in immigrant-induced job creation and productivity growth. The first set of results characterizes the relationship between immigrant worker inflows, establishment presence, and immigrant absorption in local industries. Over the three decades spanning 1980 through 2010, I find a robust, positive effect of immigrant inflows on establishment counts within local industries. This effect is driven roughly equally by establishment entry and the prevention of establishment exit, but it is the latter that is a singular driver of the relationship between immigrant inflows and job creation. Over three-fourths of net jobs created in response to immigration are accounted for by the prevention of establishment exit alone, while the expansion of continuing establishments plays a minimal role. In U.S. local industries, immigrant absorption is thus a phenomenon driven by the extensive margin of firm decision-making—particularly, the shut-down decision.

The second set of results finds that productivity is a key moderator of this shut-down decision. Using a firm-level, revenues-per-worker-based proxy as my primary measure of productivity, I study a panel of over 4.5 million establishments from 2000 through 2015. Contrary to concerns that the prevention of establishment exit may also prevent creative destruction, I find that immigrant worker inflows increase the likelihood of exit by establishments whose parent firms are in the lowest quintile of the productivity distribution. Establishments whose parent firms are outside of the lowest quintile are less likely to exit, driving the overall prevention of exit.

Motivated by these exit-specific results, I complete the picture by analyzing the impact of immigrant worker inflows on a local industry's productivity distribution as a whole. I find that the increase in establishment count induced by immigrant worker inflows between 2000 and 2015 is concentrated in the top three deciles of the productivity distribution. As with job creation, there are minimal intensive margin changes to productivity among continuing firms. These results suggest that immigrant worker inflows increase productivity in a local industry, and do so primarily through the extensive margin.

Changes to business count and composition are not built into canonical models of immigration and the labor market. In light of my empirical results, I fill this gap by developing a theoretical framework that incorporates worker nativity into a model with firm heterogeneity. I propose a mechanism that ties high productivity firms to immigrant workers: some firms pay additional fixed costs to access a technology that allows them to better utilize immigrant employees. Because these firms must pay an additional cost, they are positively selected on productivity. When immigrant exposure increases, these higher productivity, immigrant-intensive firms see larger reductions in labor costs than their lower-productivity counterparts, and competitive forces drive the lowest-productivity firms out of the market. Empirically, I show two key responses to immigrant worker inflows that are consistent with the mechanisms proposed in the model: larger labor cost reductions among establishments from higher productivity firms and a larger degree of low productivity establishment culling among native-owned compared to immigrant-owned firms. The effect of immigration on native welfare—the "immigration surplus"—in this framework hinges on changes to the composition of firms in the market. Specifically, immigrant-induced changes to the productivity composition of firms generate first-order welfare benefits. In comparison, second-order effects from wage changes that also arise in canonical, representative firm models of production are relatively muted.

The rest of this paper is organized as follows: Section 1.1 provides a brief literature review. Section 2 describes the U.S. Census data and shift-share identification approach used in subsequent analyses. Section 3 quantifies and characterizes the positive relationship between immigrant worker inflows and increased establishment presence for 1980–2010 and culminates by showing how this relationship drives immigrant absorption. Section 4 analyzes heterogeneous establishment shut-down decisions in response to immigrant worker inflows for 2000–2015, then probes the overall relationship between immigrant worker inflows and establishment presence across the productivity distribution during this time period. Section 5 synthesizes the results from Sections 3 and 4 and discusses welfare implications using a simple theoretical model. Section 6 concludes.

1.1 Related Literature

This paper is unique in presenting an exhaustive analysis of how immigrant labor supply increases impact establishment entry and exit in the U.S. and elucidating the consequences of this relationship for immigrant absorption and productivity. It builds on important literatures in labor economics, international trade, and entrepreneurship. Here, I briefly summarize some of the most closely related works.

Two related papers contain results regarding the effect of immigrant workers on establishment counts in the U.S., as in Sections 3.1.1 and 3.1.2. The most closely related work is Orrenius et al. (2020), who also find that immigrant inflows increase establishment presence and reduce establishment exit in a panel dataset of 160 U.S. Core-Based Statistical Areas for the period 1997 through 2013. Olney (2013) also presents evidence that "low-skilled" immigrants generate increased establishment presence in the 30 largest U.S. metropolitan areas using a yearly panel that covers 1998 through 2008. Relative to these works, this paper is focused not just on the effect of immigration on establishment presence, but also on the consequences of this relationship for job creation and productivity. It also introduces a longer time horizon, different identification strategy, and both a finer-grained level of study and more comprehensive coverage of the U.S. economy.

Outside of the U.S., Beerli et al. (2021) study Switzerland's abolition of restrictions on cross-border commuters and find that the availability of "higher-skilled" foreign-born workers leads to an increase in establishment presence in areas most affected by the policy. Relative to Beerli et al. (2021), this paper focuses on the long-run impacts of several immigration shocks identified by the shift-share IV approach—as opposed to a single reform—and studies an immigrant-induced labor supply increase that is majority "lower-skilled." In addition, it finds that business entry and exit are primary, rather than ancillary, mechanisms through which immigration affects job creation and productivity. This contrast can be seen in this paper's focus on the prevention of establishment exit compared to null findings for establishment exit in Beerli et al. (2021). Altindag et al. (2020) also find that the influx of Syrian refugees spurred an increase in firm presence in Turkey between 2006 and 2015. Unlike their work, this paper focuses on several long-run shocks that affect establishments in the formal economy. Furthermore, the majority of immigrant workers who generate the effects found in this paper were likely brought in through family reunification rather than as refugees.

A deep literature studies the questions of whether and how immigrants are absorbed into local economies, questions most directly addressed in Section 3.2. Motivated by Lewis (2005) and Lewis (2012), this paper is specifically focused on how labor demand responds to immigrant inflows, as opposed to other important immigrant absorbing mechanisms—including changes in output mix (e.g., Gonzalez and Ortega, 2011; Burstein et al., 2020), changes in consumer demand (e.g., Hong and McLaren, 2015), and changes to labor supply (e.g., Monras, 2020).[10] Dustmann and Glitz (2015) were the first to broach firm entry and exit as important immigrant absorption mechanisms, finding that it accounts for 15 percent of immigrant absorption in Germany's tradable sector. Their decomposition exercise motivates the decomposition presented in Section 3.2.

A growing body of work relates to the findings in Section 4 and its implications for immigrant-induced changes to productivity. Mitaritonna et al. (2017) study the effect of immigration on establishment productivity in France and find that immigrant-induced increases in productivity largely occur among continuing establishments. When placed alongside Beerli et al. (2021), the differences between Mitaritonna et al. (2017) and the results found in Section 4 broach the possibility that the U.S. is unique in the extent to which the exit margin drives the effects of immigration. Peri (2012) analyzes a decadal panel of U.S. states and finds that immigrant inflows into the workforce generate increases in total factor productivity, mediated by task reallocation and increased specialization by native workers. The empirical results presented in this paper suggest that some of these responses occur through reallocations that are enabled by establishment entry and exit.

Other studies on U.S. local labor markets indirectly comport with the finding in Section 4 that immigrant worker inflows lead to establishment-level creative destruction. Burchardi et al. (2020) analyze the same study period as this paper and find that immigration leads to increased innovation (as measured by patents), job creation, job destruction, job growth skewness, and wages at the U.S. county level. That job destruction increases in tandem with job creation—indicating an increase in dynamism—aligns with the stratified exit responses found in Section 4. Further evidence consistent with these dynamics comes from recent work by Ayromloo et al. (2020), who

[10] Important recent works comparing different channels of adjustment include Dustmann and Glitz (2015) and Monras (2021).

find that state enforcement of e-Verify laws leads to exits by larger (likely more productive) firms. By studying an establishment level panel and incorporating plausible proxies for productivity, this paper is able to make advances to this literature by showing that business entry and exit are key mechanisms underlying the link between immigration and economic dynamism.

The empirical results in this paper are also related to but set apart from recent literature on entrepreneurship. Relative to work on the link between declining population growth and declining entrepreneurship (Hopenhayn et al., 2018; Karahan et al., 2019), the results of this paper imply that immigrant inflows are distinct from general population growth, both in the magnitude of their impact on establishment presence and in the central importance of exit prevention. This distinction aside, results in Section 3.1.2 do imply that immigration can help alleviate the declining start-up rate. Findings in Sections 3.1.1 and 3.1.3 comport with recent findings on the importance of immigrant entrepreneurship to the U.S. economy (Kerr and Kerr, 2016, 2018; Azoulay et al., 2020). However, the group of workers who constitute the immigrant inflows of interest in this paper is heavily tilted towards immigrant employees, and the centrality of exit prevention indicates that these employees are the main drivers of the dynamics found here.

Section 5 develops a theoretical model with worker heterogeneity across nativity and firm heterogeneity in both total factor productivity and input use. These heterogeneities are motivated by Bustos (2011)—who introduces endogenous technological change to the Melitz framework by allowing a subset of firms to pay a higher fixed cost to access a better production technology. Tying immigrant workers to higher productivity firms introduces a new channel through which immigrants increase productivity and lower prices: by shifting firm productivity distribution rightward. Immigrant-induced price decreases that increase consumer welfare have been studied by Cortes (2008) (through a wage reduction channel) and both di Giovanni et al. (2014) and Hong and McLaren (2015) (through an increased variety channel). The productivity distribution channel is a novel addition to this theoretical literature. Independently, Brinatti and Morales (2021) also develop a model that incorporates worker heterogeneity by nativity into a model of firm heterogeneity and find that this combination increases the immigration surplus. Unlike their work, this paper is focused on the extensive margin and the firm distribution rather than the decisions of continuing firms on the intensive margin.

2 Data and Identification

2.1 Data

The analyses presented in Sections 3 and 4 are facilitated by access to confidential data from the U.S. Census Bureau's Longitudinal Business Database (LBD), which is constructed from administrative tax records for each U.S. non-farm, employee-hiring, private-sector establishment. Establishments are assigned unique, consistent identifiers that can be linked over time to create a true panel. The LBD also contains unique firm identifiers, which allows me to aggregate establishments to their parent firms. A majority of establishments can be linked to firm-level revenue information from the Census Bureau's BRFIRM_REV dataset starting in 1997 (see Haltiwanger et al., 2019). For a representative sample in 2007, these firm identifiers can also be linked to firm ownership nativity information from the Survey of Business Owners (SBO).

In order to study the effect of immigrant presence on outcomes constructed from the LBD, I also exploit restricted-access U.S. Census Bureau demographic data from the 1980, 1990, and 2000 Long-Form Decennial Censuses and the 2005 through 2019 American Community Surveys (ACS). The demographic data allow for unusually precise measures of immigrant inflows, not just into geographies, but into relatively detailed industry groups within U.S. counties and by country of origin. These elements are important to the identification strategy presented in Section 2.2. I limit the sample to employed individuals (both self-employed and employees) so that I can assign workers an industry group.

The analyses presented in Section 3 and Section 4 are based on immigrant exposure in commuting zone-industry group pairings—local industries—over time.[11] I study the 722 commuting zones in the contiguous United States and the 41 industry groups seen in Table A1. This results in coverage of 29,602 local industries per time period in each estimated model. Appendix Section A provides additional details on the data and sample construction, including Section A.4 on industry group construction.

2.2 Identification Strategy

To facilitate discussion of the identification strategy, I first present the primary specification used in Section 3:[12]

$$\Delta y_{gkt} = \beta\left(\Delta I_{gkt}\right) + \Gamma X_{gkt} + \alpha_{gt} + \alpha_{r(g),kt} + \varepsilon_{gkt} \tag{2.1}$$

where g indexes a commuting zone, $r(g)$ indexes the Census Region that contains commuting zone g, k indexes an industry group, t indexes a year, and the Δ operator represents a ten-year change within a local industry (change within gk).[13] Δy_{gkt} is an outcome related to establishment entry and exit dynamics in local industry gk over the decade that ends in year t. The independent variable of interest, ΔI_{gkt}, is the change in immigrant worker stock in gk between year $t-10$ and t, divided by the start-of-decade workforce in gk—a relative immigration shock to labor supply.[14] X_{gkt} is a vector of control variables that are described in more detail below. First differencing removes the influence of time-invariant local industry confounders.

Even with the rich fixed effect structure contained in Equation (2.1), there are endogeneity concerns regarding immigrant industry choices within geographies and geographical choices within industry. Immigrant employees in a given commuting zone may choose to work in its fastest-

[11] Commuting zones are groupings of counties meant to mimic local labor markets. Crosswalks from counties to commuting zones are downloaded from David Dorn, as used in Autor and Dorn (2013).

[12] A modified version of Equation (2.1), with Δ representing 15 year changes, is also used in Section 4.3, and an analogous specification brings Equation (2.1) to the establishment level for Sections 4.1, 4.2, and 5.5.

[13] From this point forward, region will refer to a Census Region: East, West, South, or Midwest.

[14] "Relative" refers to the fact that the immigrant stock is divided by start-of-period workforce size.

growing industries, generating upward-biased estimates of β. A similar bias would arise if immigrant workers were more adept than native workers at moving to areas where their industries are demanding more labor—a within-industry extension of the results found in Cadena and Kovak (2016). Meanwhile, if immigrant entrepreneurs are attracted to geographies where they face less competition, or if immigrant employees are linked to large firms in more concentrated markets, OLS estimates of β may be downward-biased when the outcome is related to establishment counts. In short, even in a relatively saturated model, isolating exogenous variation that pushes immigrants into local industries substantially strengthens our case for a causal interpretation of β. To this end, I next turn to a shift-share instrumental variables approach.

2.2.1 Emigrants IV

A standard shift-share instrument for the relative stock of immigrants in local industry gk and year t takes the following form:

$$z_{gkt}^{\text{Standard}} \equiv \frac{1}{E_{gk,1980}} \sum_{o} \pi_{og,1980} \times \mathcal{I}_{okt}$$

where $\pi_{og,1980}$ is the share of origin country o's stock of 1980 U.S. immigrants that was located in commuting zone g, $\mathcal{I}_{okt}$ is the total stock of year t U.S. immigrants from origin country o working in industry group k, and $E_{gk,1980}$ is the 1980 workforce in local industry gk. In this standard setup, $\Delta z_{gkt}^{\text{Standard}}$ would serve as an instrument for ΔI_{gkt} in Equation (2.1), with shift component $\Delta\mathcal{I}_{okt}$ representing net worker inflows into industry group k from origin country o.

Previous literature has implicitly recognized the value of an exogenous shift in a shift-share instrument. In a well-known, recent example, Autor et al. (2013) instrument for Chinese import exposure in U.S. local labor markets by interacting measures of initial industry concentration in commuting zones (shares) with Chinese, industry-level *export* growth to eight advanced, *non-U.S.* economies (shifts). Similar exogenous-shift strategies have been employed in the immigration literature. Llull (2017) for example, uses a suite of origin-specific migration push factors—conflict, natural disasters, changes in per capita income, and changes to political regimes—as shifts in shift-share instruments. In both Autor et al. (2013) and Llull (2017), the use of shifts that originate from abroad is an attempt to purge a shift-share instrument of latent economic trends specific to any of the local labor markets being studied.

Borusyak et al. (2020) and Jaeger et al. (2018) provide some explicit theoretical backing for the use of these plausibly exogenous shifts. Borusyak et al. (2020) show that when shifts represent a set of relatively disbursed and uncorrelated shocks across origin countries, their quasi-random assignment conditional on shares can overcome endogeneity in the share component of the instrument.[15,16] Meanwhile, Jaeger et al. (2018) specifically focus on shift-share instruments for im-

[15]Since this identification strategy is "conditional on shares," $\sum_o \frac{\pi_{og,1980}}{E_{gk,1980}}$, interacted with year fixed effects, is always included in X_{gkt}.

[16]This identifying assumption contrasts with Goldsmith-Pinkham et al. (2020), who focus on exogeneity in the share component as a sufficient condition for instrument validity.

migration and show that aggregate immigrant inflows into the U.S. are highly serially correlated in their origin country composition. This generates a potentially severe bias in which panel regressions can confound short- and long-run responses to immigration: $\Delta z_{gkt}^{\text{Standard}}$ can often be just as—if not more—correlated with $\Delta I_{gk,t-10}$ than it is with ΔI_{gkt}. Thus, to the extent that they are more plausibly exogenous and less serially correlated at the origin-country-level than aggregate inflows, alternate push factors that replace ΔI_{okt} are also less likely to generate biased coefficient estimates $\hat{\beta}$.

I thus construct a shift-share instrument in the vein of Llull (2017) and Autor et al. (2013). First, I note that:

$$z_{gkt}^{\text{Standard}} = \frac{1}{E_{gk,1980}} \sum_{o} \pi_{og,1980} \times \underbrace{\frac{\mathcal{I}_{okt}}{\mathcal{I}_{ot}}}_{\equiv \rho_{okt}} \times \mathcal{I}_{ot} \tag{2.2}$$

where ρ_{okt} is the share of origin country o's stock of year t U.S. immigrants that are working in industry group k, and $\mathcal{I}_{ot}$ is origin country o's stock of year t U.S. immigrants. Next, utilizing unique bilateral emigrant stock data from the German Institute for Employment Research's (IAB) Brain-Drain Data (Brücker et al., 2013) and the United Nations Population Division's (UNPD) International Migration Stock 2019, I replace the final $\mathcal{I}_{ot}$ in (2.2) with $\mathcal{M}_{ot}^{\text{non-U.S.}}$, the count of emigrants from origin country o living in 18 OECD member nations *other than* the United States.[17]

By replacing $\mathcal{I}_{ot}$ with $\mathcal{M}_{ot}^{\text{non-U.S.}}$, I retain the full suite of factors that are *pushing* individuals in origin country o to emigrate—both to the U.S. and to other destinations—while discarding any factors that are specifically *pulling* immigrants into the U.S.—including and most importantly, local-industry-specific labor demand.[18] Just as Autor et al. (2013) claim that Chinese exports to non-U.S. countries reflect increases in Chinese export productivity rather than product demand in the U.S., I claim that non-U.S. emigrant outflows are much more likely to reflect migration push factors in origin countries rather than demand in specific U.S. labor markets.

Finally, exploiting the detailed information on industry and origin country contained in the U.S. Census Bureau's demographic data, I am able to replace ρ_{okt} in (2.2) with $\rho_{okt,-r(g)}$: the year t share of immigrants from origin country o that work in industry k, in all areas *other than the region $r(g)$ that contains commuting zone g*. Allocating emigrants into industry groups using $\rho_{okt,-r(g)}$ takes advantage of the fact that immigrants from particular countries tend to specialize in certain industries due to comparative advantage, but once again discards demand factors from the specific local industry being studied by leaving out region $r(g)$. $\rho_{okt,-r(g)}$ performs the role of disbursing aggregate non-U.S. emigrant stocks into industry groups in the way that $\pi_{og,1980}$ disburses them into commuting zones.

[17] See Appendix Section A.3 for more details on the emigration data.

[18] Orefice and Peri (2020) independently use a similar "leave-country-out" instrumentation strategy for immigration to French regions.

The result is a new instrumental variable for relative immigrant presence in a local industry:

$$z_{gkt}^{\text{Emigrants}} \equiv \frac{1}{E_{gk,1980}} \sum_{o} \pi_{og,1980} \times \rho_{okt,-r(g)} \times \mathcal{M}_{ot}^{\text{non-U.S.}} \tag{2.3}$$

$z_{gkt}^{\text{Emigrants}}$ predicts the number of immigrants residing in a given commuting zone g based on network-induced locational preference, working in industry k due to country-specific comparative advantage, and pushed into the U.S. by factors that also pushed their compatriots to emigrate to non-U.S. destinations. In Sections 3 and 4.3, $\Delta z_{gkt}^{\text{Emigrants}}$ will serve as the instrumental variable for ΔI_{gkt}, and in Sections 4 and 5.5, $z_{gkt}^{\text{Emigrants}}$ will serve as an instrumental variable for I_{gkt}.

$\Delta z_{gkt}^{\text{Emigrants}}$ is a valid instrument for ΔI_{gkt} if shifts $\Delta\left[\rho_{okt,-r(g)} \times \mathcal{M}_{ot}^{\text{non-U.S.}}\right]$ are conditionally uncorrelated with unobserved factors in the local industries whose commuting zones have the highest shares $\pi_{og,1980}$.[19] Violations can occur when labor demand shocks pull immigrants into local industries in areas where their compatriots are already heavily concentrated. Appendix Section B.6.3 examines the identifying variation in the construction sector during the housing bubble and shows how the instrumental variable helps eliminate the influence of endogenous immigrant inflows during this time period. Section 2.2.3 presents more systematic evidence in the same vein.

2.2.2 *Fixed Effects*

With the instrument fully detailed, the utility of each set of fixed effects in Equation (2.1) becomes more clear. For the decade ending in year t, α_{gt} removes any effects immigrant inflows have at the commuting zone level as a whole. Under the premise that immigrants do not solely demand goods in the industry in which they work, α_{gt} then insulates β from being primarily identified by changes in consumption patterns that can result from immigration. This premise is strengthened by the fact that we compare across 40 industry groups within a commuting zone—a level of detail allowed for by the granularity in each Census Bureau data source.[20] Removing the impact of consumer demand helps align the identification strategy with the goal of this paper, which is to specifically examine how labor demand accommodates a shock to labor supply.[21] Appendix Section D.2 presents a simplified version of the model presented in Section 5 that directly shows how the inclusion of α_{gt} eliminates the influence of consumer demand on estimates of β.

$\alpha_{r(g),kt}$ plays an important role in ensuring instrument validity. $\rho_{okt,-r(g)}$ allocates immigrants from origin country o into industry group k based on national level trends of industry choice for origin country o immigrants, excluding region $r(g)$. In the absence of $\alpha_{r(g),kt}$, national level shocks

[19] Models in Section 4 contain establishment fixed effects, which take the place of the gk-level differencing seen in Equation (2.1). So, one can view the identifying assumption in the same way.

[20] For example, an inflow of immigrants into the "Hospitals" industry group can generate an increase in economic activity in other nontradable industry groups because the new immigrant workers in the "Hospitals" industry group also consume goods and services locally. By including α_{gt} and thus inducing comparison across industry groups within a given commuting zone and decade, β measures the immigrant-induced increase in economic activity in the "Hospitals" industry group above and beyond what other industry groups experienced due to this consumer demand effect.

[21] As opposed to a simultaneous shock to labor supply and consumer demand.

to industry k would naturally allocate all workers towards industry group k, regardless of origin. Instead, the inclusion of $\alpha_{r(g),kt}$ effectively limits a local industry's comparison group to its counterparts within the same region. $\rho_{okt,-r(g)}$ can then be thought of as a measure of comparative advantage that is measured externally, as if coming from the origin country o itself and only impacting local industries within $r(g)$ differently through $\pi_{og,1980}$. $\Delta z_{gkt}^{\text{Emigrants}}$ is therefore credibly free from contamination by demand shocks in commuting zone g and sourced from origin-specific comparative advantages through $\rho_{okt,-r(g)}$ when $\alpha_{r(g),kt}$ is included.

2.2.3 *Emigrants IV Diagnostics*

To more systematically evaluate the plausibility of the exclusion restriction for $\Delta z_{gkt}^{\text{Emigrants}}$, I conduct several balance tests of the type advocated in Borusyak et al. (2020), estimating

$$\Delta y_{gk,pt}^{\text{Std.}} = \beta_{\text{Balance}} \left(\Delta z_{gkt}^{\text{Emigrants}}\right) + \alpha_{gt} + \alpha_{r(g),kt} + \varepsilon_{gkt} \tag{2.4}$$

for the decades spanning 1980–2010, where $\Delta y_{gk,pt}^{\text{Std.}}$ are "balance outcomes:" pt indicates that they are measured prior to decade start, and "Std." indicates that all outcomes are standardized for comparability. The cleanest test of instrument balance uses 1970s (1970–1980) employment growth from the Decennial Census as a balance outcome—essentially a pre-trends test at the local industry level. I also include similar pre-trends balance outcomes from LBD-measured employment and establishment count growth, but these can only be calculated for the time period 1976-1980.[22] For comparison, I also include the standardized versions of these outcomes *during the study period*. I term these "true outcomes." Finally, I also include the control variables that are included in my primary specifications as balance outcomes.[23]

The results of these exercises are presented in Figure 1. Each of the pre-period balance outcomes fail to show evidence of pre-existing local industry growth, while the true outcomes are meaningfully larger in magnitude and statistically significant. Among control variables, the lack of correlation between $\Delta z_{gkt}^{\text{Emigrants}}$ and a Bartik labor demand control bolsters the case that the instrumental variable isolates supply pushes from demand pull factors. There is a significant, negative correlation between start-of-decade college share in a local industry and $\Delta z_{gkt}^{\text{Emigrants}}$. While it therefore appears important to control for start-of-decade college share, it is notable that we expect this variable to positively correlate with outcomes like establishment growth. The same is true for start-of-decade self-employment share. Finally, remaining concerns should be alleviated by the stability of parameter estimates to the inclusion or exclusion of these control variables (see Table 2).

I conduct additional diagnostics around $\Delta z_{gkt}^{\text{Emigrants}}$ as well. I apply the double-instrumentation procedure advocated by Jaeger et al. (2018) in order to avoid confounding of short- and long-run re-

[22]As described below, all growth rates are calculated as Davis-Haltiwanger-Schuh growth rates: Growth Rate in x between $t = 0$ and $t = 1 = \frac{x_1 - x_0}{\left(\frac{x_1 + x_0}{2}\right)}$.

[23]See Appendix Section A.5 for details on control variables and their construction.

sponses in Appendix Section B.6.1 and do not find strong evidence that this bias drives results. Beyond instrument validity, Appendix Section B.6.2 addresses inference concerns broached by Adao et al. (2019) with a simulation exercise, and finds that they are not operative for $\Delta z_{gkt}^{\text{Emigrants}}$. Appendix Section B.6.4 shows that $\Delta z_{gkt}^{\text{Emigrants}}$ performs substantially better than $\Delta z_{gkt}^{\text{Standard}}$ in terms of balance and double-instrumentation.

All told, the emigrants-based instrument holds up to a battery of checks. $\Delta z_{gkt}^{\text{Emigrants}}$ therefore becomes the instrument of choice for Sections 3 and 4.3, while $z_{gkt}^{\text{Emigrants}}$ becomes the instrument of choice for Sections 4.1, 4.2, and 5.5.

2.2.4 *Education Level of Immigrant Inflows*

Contextualizing results generated by the identification strategy outlined above requires a more complete understanding of the immigrant inflows represented by ΔI_{gkt}, particularly when it is instrumented for by $\Delta z_{gkt}^{\text{Emigrants}}$. To take one common motivation, consider the seminal Borjas (1999) "immigration surplus" insight: when there are extant skill differences across native and foreign-born workers, the average set of native skills becomes more scarce when immigrant workers flow into the economy, generating average wage gains for native workers.

Because of how they proxy for differences in labor market skill, differences in educational attainment across immigrants and natives are a crucial channel through which a connection between immigration and economic activity can arise in a manner that does not just reflect general population growth.[24] Figure 2 benchmarks the differences in educational attainment associated with immigrant inflows relative to the receiving native workforces in the local industries studied throughout this paper. The bars reflecting immigrant inflows are obtained by replacing Δy_{gkt} with $\Delta I_{gkt}^{\text{Educ. Level}}$ in Equation (2.1), where $\Delta I_{gkt}^{\text{Educ. Level}}$ represents the change in the number of immigrants in a given local industry of a given educational attainment level.[25] Exploiting the adding-up property of linear regression, estimating Equation (2.1) for mutually exclusive and exhaustive educational groupings decomposes how many workers of each educational attainment category are brought in by each immigrant, on average. The left bar of Figure 2 provides the average educational attainment of the receiving native workforce across decade start-years—1980, 1990, and 2000—for comparison.

Figure 2 contains two key findings. First, comparing the "IV Inflow" bar to the "Start-of-Decade Native Workforce" bar, the degree to which $\Delta z_{gkt}^{\text{Emigrants}}$-pushed immigrants change the educational composition of the workforce depends on how we classify workers based on education. On average, foreign-born workers who flow into the U.S. between 1980 and 2010 are substantially less likely to have a high-school degree than native workers but just as likely to have a college degree. If we take this paper's preferred measure, which compares "high-school equivalent" to "college equivalent

[24] Though it is far from the only: immigrants can differ on other characteristics that also imbue them with different labor market skills than natives. For example, lower-educated immigrant workers may have a comparative advantage in performing tasks that are less communication-oriented relative to lower-educated native workers (Peri and Sparber, 2009).

[25] $\sum_{\text{Educ. Level}} \Delta I_{gkt}^{\text{Educ. Level}} = \Delta I_{gkt}$.

workers" we find that $\Delta z_{gkt}^{\text{Emigrants}}$-pushed immigrants do tilt the workforce towards lower-educated workers.[26] On the education dimension, then, standard immigration surplus arguments should apply. Second, comparing the "IV Inflow" bar to the "OLS Inflow" bar reveals that $\Delta z_{gkt}^{\text{Emigrants}}$ tends to push immigrants of slightly higher educational attainment into the U.S. relative to the typical immigrant inflow during the study period. Using similar figures, Appendix Section B.6.5 provides further details on how $\Delta z_{gkt}^{\text{Emigrants}}$ compares to the average immigrant inflow in terms of class of work and origin country region.

3 Establishment Entry and Exit and Immigrant Absorption (1980–2010)

Using Equation (2.1), this section studies three decades in which there were large immigrant inflows to the U.S. and makes two novel points. First, immigrant worker inflows generate large responses on the extensive margin, both through establishment entry and exit. This effect is roughly equally driven by establishment entry and exit and persists throughout the employer size distribution. Additional heterogeneity analyses described in Appendix Section B.2 also find that this effect is larger in high-school-equivalent-hiring industries and tradable industries. Second, it is precisely through these large extensive margin responses that immigrant workers are absorbed into the local industries they enter. Strikingly, I cannot reject the null hypothesis that this absorption capacity is entirely driven by establishment entry and the prevention of establishment exit, with the latter playing a dominant role.

3.1 The Effect of Immigrant Worker Inflows on Establishment Presence

3.1.1 Quantifying the Effect

I start by estimating the effect of immigrant worker inflows on the Davis-Haltiwanger-Schuh (DHS) growth rate in local industry establishment count:

$$\Delta y_{gkt} = \frac{\Delta \text{Estabs}_{gkt}}{\text{Estabs Denom}_{gkt}} = \frac{\text{Estabs}_{gkt} - \text{Estabs}_{gk,t-10}}{\left(\frac{\text{Estabs}_{gkt} + \text{Estabs}_{gk,t-10}}{2}\right)}$$

where Estabs_{gkt} is the count of establishments operating in local industry gk and time t. An establishment is defined as operating if it has both positive payroll and employment. Use of DHS growth rates allows me to retain some of the useful properties of log changes while also retaining the ability to decompose the numerator $\text{Estabs}_{gkt} - \text{Estabs}_{gk,t-10}$ into component parts in subsequent analyses.[27]

[26] High-school equivalent workers are those with a high school degree or less plus half of those with some college. College equivalent workers are half of those with some college plus those with a college degree or more. The relevant comparison, then, can be seen in the middle of the purple bar across columns.

[27] Appendix Section D.2 presents a simplified version of the model presented in Section 5 that motivates the use of log changes (or the DHS growth rates that closely align with them).

Table 1 displays first stage, OLS, and IV results from estimating Equation (2.1) using this outcome variable. I find a strong, positive effect of immigrant worker presence on establishment presence that is larger when corrected for endogeneity. A one percent shock to a local industry's relative labor supply induced by immigration generates a 0.75 percent increase in the establishment count, on net (Column 3). This corresponds to four additional establishments generated per 100 immigrant workers that flow into a local industry.[28]

Table 2 demonstrates the stability and robustness of this estimate to different control and fixed effects sets and also provides some context around its magnitude. Moving from Column 1 to Column 2 indicates the utility of including $\alpha_{r(g),k,t}$ and α_{gt} relative to a traditional first differences approach (just including α_t). Nearly half of the first difference estimate disappears when we compare across industry groups within a commuting zone, relative to when we simply compare across local industries. Columns 2 and 3 show that results are invariant to the inclusion of control variables, stability that bolsters the case for instrument validity.[29]

Column 4 of Table 2 includes an endogenous control for native inflows per initial worker. That Column 3 and Column 4 deliver near-identical results imply three important points. First, immigrant workers appear to increase establishment presence substantially more than native workers. A one percent relative labor supply shock that comes from immigrant workers is roughly seven times more impactful on establishment presence than a corresponding relative labor supply shock that comes from native workers. While immigration is a form of more general population growth, these results suggest that immigrant workers' effects on business dynamics are distinct. Second, the effect of immigration on establishment presence does not appear to be attenuated by native mobility responses. Third, that instrument strength does not wane with the inclusion of this control strengthens the argument that it is operating through immigrant inflows specifically and not workforce growth generally.

Finally, Column 5 shows that the DHS growth rate in establishment count outcome variable delivers near-identical results to using the change in log establishment count (comparing Column 5 to Column 3). This demonstrates the ability of DHS growth rate variables to approximate changes in log counts well while still allowing for important decompositions of overall changes, which I present next.

3.1.2 *Characterizing the Effect: Entry and Exit*

Figure 3 decomposes the change in the count of establishments into its flow components. Letting e index an establishment, $\mathcal{E}_{gkt}$ denote the set of establishments that entered local industry gk between $t-10$ and t (entrants), and $\mathcal{X}_{gkt}$ denote the set of establishments that were alive in $t-10$ that

[28] This correspondence is obtained from using $\Delta y_{gkt} = \frac{\text{Estabs}_{gkt} - \text{Estabs}_{gk,t-10}}{\text{Workers}_{gk,t-10}}$ to estimate Equation (2.1). In this case Δy_{gkt} and ΔI_{gkt} share a denominator, so β is interpreted as "establishments per immigrant." $\hat{\beta}\,(SE) = 0.0413\,(0.0061)$.

[29] Note that Column 3 of Table 2 is identical to Column 3 of Table 1.

were not operating in t (exiters):[30]

$$\underbrace{\frac{\Delta \text{Estabs}_{gkt}}{\text{Estabs Denom}_{gkt}}}_{\text{Outcome: Total}} = \underbrace{-\frac{\sum_e \mathbb{1}\left\{e \in \mathcal{X}_{gkt}\right\}}{\text{Estabs Denom}_{gkt}}}_{\text{Outcome: Exit Prevention}} + \underbrace{\frac{\sum_e \mathbb{1}\left\{e \in \mathcal{E}_{gkt}\right\}]}{\text{Estabs Denom}_{gkt}}}_{\text{Outcome: Entry}}$$

This decomposition is only available due to longitudinal linkages provided in the LBD.

Figure 3 plots the results of estimating Equation (2.1) using these disaggregated outcomes, such that the OLS bar on the left totals 0.3972 (Column 2 of Table 1) and the IV bar on the right totals 0.7509 (Column 3 of Table 1) and Column 3 of Table 2. Focusing on the IV results, both establishment entry and the prevention of establishment exit appear to play important roles in generating the overall increase in establishment presence that arises in response to immigrant worker inflows, with entry accounting for 43 percent and exit prevention accounting for 57 percent.

These results comport with and extend important literature on immigration and business dynamics. Exit prevention is also found in Orrenius et al. (2020). Its quantitative importance also makes it unlikely that the overall impact of immigrant worker inflows on establishment presence is driven solely by *new* immigrant entrepreneurship—that is, entry by immigrant-owned businesses.[31] Back-of-the-envelope calculations presented in Appendix Section B.7.1 show that new immigrant entrepreneurship could nonetheless be responsible for up to 62 percent of the establishment *entry* found in Figure 3. This is consistent with a growing body of literature on the importance of immigrant entrepreneurs to establishment entry (e.g., Azoulay et al., 2020; Kerr and Kerr, 2016, 2018).[32]

3.1.3 Characterizing the Effect: Establishment Size

It is important not only to understand how the count of establishments changes in response to immigrant worker inflows, but how this count is reflected in the size distribution. Increased establishment presence, for example, may not impart substantial economic activity on a local industry if these establishments do not employ many workers.[33]

To address these concerns, I estimate Equation (2.1) using DHS growth rates in establishment

[30] For the remainder of the paper, I will use "exit" and "not in operation" synonymously. Note that this means that "exit" is not necessarily an absorbing state in this paper. Empirically, "exit" and "not in operation" coincide in the large majority of cases.

[31] While it is possible that input-output linkages can generate a link between new immigrant entrepreneurship and the prevention of existing establishment exit, the arrival of immigrant employees has more clear and direct direct channels to preventing exit, including labor cost reductions and production complementarities. Furthermore, cross-establishment linkages often occur across industries, and these effects are netted out by the fixed effect structure in Equation (2.1).

[32] Furthermore, *continuing* immigrant-owned businesses also play a role in the prevention of establishment exit. As seen in Section 5.5.2, there is a significant decrease in exit probability among immigrant-owned firms in response to immigrant worker inflows, on average.

[33] These establishments would be termed "subsistence" in the parlance of Schoar (2010)

counts within employee size bin as outcomes:

$$\Delta y_{gkt} = \frac{\Delta \text{Estabs}_{gkt}^{\text{Size Bin}}}{\left(\frac{\text{Estabs}_{gkt}^{\text{Size Bin}} + \text{Estabs}_{gk,t-10}^{\text{Size Bin}}}{2}\right)}$$

I calculate growth rates within bin because size distribution is heavily skewed right.

Figure 4 finds a robust, large effect of immigrant inflows on increased establishment presence throughout the size distribution, and one that is largest at both tails. This finding aligns with results in Azoulay et al. (2020), who find that immigrant entrepreneurs own firms throughout the size distribution. However, the results here show that the *overall* effect of immigrant worker inflows—including employees—has a similar effect on the size distribution. That this effect persists across small and large establishments alike suggests that increased establishment presence may play a critical role in mediating immigrant-induced net job creation. I test this hypothesis directly, next.

3.2 The Role of Firm Entry and Exit in Immigrant Absorption

Motivated by the large effects of immigrant worker inflows on establishment entry and exit throughout the size distribution found in Section 3.1, this section asks whether and how immigrant workers are absorbed into labor industries—or, put differently, whether and how a local industry creates and sustains enough jobs to keep pace with immigrant inflows.

Letting C_{gkt} denote the set of establishments that were operating in industry gk in $t-10$ and are still operating in local industry gk at time t, I utilize the following decomposition of employment, enabled by the longitudinal structure of the LBD:

$$\underbrace{\frac{\Delta \text{Employment}_{gkt}}{\text{Employment Denom}_{gkt}}}_{\text{Outcome: Total}} = -\overbrace{\frac{\sum_{e \in \mathcal{X}_{gkt}} \text{Employment}_{egk,t-10}}{\text{Employment Denom}_{gkt}}}^{\text{Outcome: Contribution of Exit Prevention}} + \overbrace{\frac{\sum_{e \in \mathcal{E}_{gkt}} \text{Employment}_{egkt}}{\text{Employment Denom}_{gkt}}}^{\text{Outcome: Contribution of Entrants}}$$
$$+ \underbrace{\frac{\sum_{e \in C_{gkt}} \Delta \text{Employment}_{egkt}}{\text{Employment Denom}_{gkt}}}_{\text{Outcome: Contribution of Continuers}} + \text{Residual}$$

where Employment $\text{Denom}_{gkt} = \left(\frac{\text{Employment}_{gkt} + \text{Employment}_{gk,t-10}}{2}\right)$. The first term represents job loss from establishment exits between $t-10$ and t in local industry gk, the second term represents gross job creation at establishments that were born between $t-10$ and t, and the third term represents employment growth at establishments that were alive in both $t-10$ and t and in operation in gk.

Figure 5 is the clearest illustration that the extensive margin drives labor demand responses to immigration. The IV-estimated decomposition in the right bar shows that entry and exit play the dominant role in immigrant absorption, with the prevention of establishment exit on its own accounting for more than 80 percent of immigrant-labor-supply-induced net job creation. Meanwhile, Table A4 shows that I cannot reject the null hypothesis that continuing establishments play no role in immigrant absorption. At the upper limit of its 95% confidence interval, continuing establishments account for 38 percent of immigrant-induced job creation; meanwhile, at the lower limit of its confidence interval, exit prevention accounts for 48 percent. Imprecise estimates preclude strong inference regarding the role of entry, though point estimates indicate that it plays a larger role than continuing establishment expansion as well. In sum, immigrant absorption appears to revolve around establishment entry and exit—particularly exit prevention—in the context of U.S. local industries.

The IV-estimated "Total" coefficient (black dot on the right of Figure 5) indicates that a one percent shock to relative labor supply induced by immigration generates a 0.43 percent increase in the employment growth rate. This translates to roughly 0.53 LBD jobs created, on net, per immigrant worker.[34] There is no evidence for spillover job creation, and some indirect evidence for native industry-, geography-, or industry-and-geography-switching in response. This consistent with other studies that examine immigrant-induced labor supply shocks and particularly unsurprising given that α_{gt} controls away much of the influence of immigrant consumer demand and that there are 41 industry groups studied within each commuting zone.[35]

The OLS "Total" estimate on the left of Figure 5 is substantially larger than the IV estimate on the right (by more than 50 percent). This accords with the notion that immigrants are attracted to local industries with increasing labor demand and with the claim that $\Delta z_{gkt}^{\text{Emigrants}}$ corrects for some of this endogeneity. It also adds context to results from Figure 3, which found that IV estimates of immigrant worker inflows on establishment presence were higher than OLS estimates. One possible explanation that squares these results is that immigrant entrepreneurs are attracted to areas with lower net firm entry if these areas feature less potential competition.[36] However, as noted above, immigrant entrepreneurship is not likely to be the sole driver of the effects found in here. A complementary explanation is that areas and industries with higher employment concentration among larger firms are more likely to attract immigrants. Comparing components across the OLS-estimated and IV-estimated results also suggest interesting conclusions regarding immigrant locational and industry choices: immigrants appear attracted to geographies and industries

[34]This correspondence is obtained from using $\Delta y_{gkt} = \frac{\text{Employment}_{gkt} - \text{Employment}_{gk,t-10}}{\text{Workers}_{gk,t-10}}$ to estimate Equation (2.1). In this case Δy_{gkt} and ΔI_{gkt} share a denominator, so β is interpreted as "jobs per immigrant." $\hat{\beta}\,(SE) = 0.5266\,(0.1525)$.

[35]A bevy of previous literature has documented native mobility across occupations in response to immigrant inflows. Burstein et al. (2020), for example, document substantial native displacement in non-tradable industries. Monras (2021) documents that the "low-skilled" workforce in Miami increases by 0.6 for every immigrant brought in by the Mariel Boatlift. It would stand to reason that the number here should be slightly lower, given that switches within commuting zone but across industry also reduce the number of jobs counted and that ΔI_{gkt} includes self-employed individuals and employees in the public sector. Hong and McLaren (2015) find substantial spillover job creation, but at the commuting-zone-wide level. Moreover, they hypothesize that this spillover effect is the result of consumer demand, which I attempt to control away here.

[36]See., e.g., Ottinger (2020) for historical evidence that immigrants migrated to U.S. counties that were less specialized in the industries in which they worked.

whose establishments are growing and that are experiencing establishment entry. The stronger negative effect on exit prevention in the IV-estimated results may also indicate that immigrants are attracted to more dynamic local industries, in which productive turnover is taking place.

4 Immigrant Workers, Establishment Exit, and Productivity (2000–2015)

The particular importance of exit prevention found in Section 3 leads to potential concerns regarding business dynamism; specifically, that it reflects stunted creative destruction and therefore decreased productivity. In this section, I thus analyze the productivity consequences of immigrant worker inflows. In Section 4.2, I stratify an establishment-level analysis of exit by proxies for initial parent firm total factor productivity. Contrary to the above concerns, I find that immigrant inflows tend to benefit more productive firms by reducing shut-down probabilities of their establishments, while culling lower productivity establishments from low productivity firms from the market. I then explore how immigrant worker inflows change a local industry's firm productivity distribution as a whole. I find that the overall increase in establishment presence induced by immigrant worker inflows is concentrated in the top three deciles of firm productivity, implying a novel, direct link between immigrant worker inflows and productivity.

The analyses in these sections utilize a panel of establishments that were operating in the U.S. as of 2000. The breadth of the LBD ensures that this panel contains near-universal coverage of the U.S. private sector. It follows these establishments for every five years through 2015. These timing restrictions are necessitated by three constraints: 1) sub-state (e.g., commuting zone), annual estimates of the foreign-born workforce are only available in 2000 (Decennial Census Long-Form), and 2005 onward (ACS); 2) the instrumental variable $z_{gkt}^{\text{Emigrants}}$ relies on emigrant counts that are only available in the UNPD data at a five-year frequency; and 3) firm-level revenue information from the BRFIRM_REV—critical for this section—can only be linked to LBD establishments starting in 1998. Of the 6.18 million establishments in the overall panel, 4.74 million can be linked to the BRFIRM_REV.

4.1 Overall Establishment-Level Responses to Immigrant Worker Inflows

As a starting point, I employ the following specification, analogous to (2.1), but broken down to the establishment level:[37]

$$y_{et} = \gamma\left(I_{gkt}\right) + \Gamma X_{gkt} + \alpha_e + \alpha_{gt} + \alpha_{r(g),kt} + \varepsilon_{et} \tag{4.1}$$

Table 3 presents the results of estimating Equation (4.1), with Panel B presenting IV estimates that use $z_{gkt}^{\text{Emigrants}}$ as the instrument for I_{gkt}. The key outcome variable is $\mathbb{1}\{\text{Not Operating}\}_{et}$, an indicator for whether or not establishment e has stopped operating at time t. This outcome is

[37] Establishment fixed effects α_e subsume commuting zone-industry group fixed effects α_{gk} and therefore the first difference seen in (2.1). Note that g and k are functions of e, but the notation $g(e)$ and $k(e)$ are omitted for ease of viewing.

compared to several intensive margin outcomes: establishment employment, establishment payroll per worker, firm-level revenues, and firm-level revenues per worker.

Panel B of Table 3 confirms a key insight found in Section 3.2 at the establishment level: extensive margin establishment responses dominate intensive margin establishment responses to immigrant worker inflows. This is true both in terms of statistical and economic significance. Even taking statistically insignificant point estimates at face value, the effect of a one percent immigration shock to relative labor supply between 2000 and 2015 generates increases to intensive margin outcomes of between 0.03 and 0.12 percent. Meanwhile, back of the envelope calculations indicate that this same shock decreases exit probability by at least 0.52 percent.[38]

4.2 The Heterogeneous Effect of Immigrant Worker Inflows on Establishment Exit

I next turn to estimating an expanded version of Equation (4.1) with $\mathbb{1}\{\text{Not Operating}\}_{et}$ as the outcome, but with the effect of I_{gkt} stratified based on where an establishment's parent firm ranks in the productivity distribution:

$$\mathbb{1}\{\text{Not Operating}\}_{et} = \sum_{d=1}^{10} \gamma_d \left(I_{gkt} \times \mathbb{1}\left\{f(e) \in d\right\}\right) + \Gamma X_{gkt} + \alpha_e + \alpha_{gt} + \alpha_{r(g),kt} + \varepsilon_{et} \tag{4.2}$$

where d indexes a decile in the productivity distribution and $f(e)$ indexes establishment e's parent firm. Based on results from Section 4.1, we expect a majority of $\gamma_d < 0$. More importantly, concerns regarding stunted creative destruction dynamics would manifest as $\gamma_d < 0$ for low d.

My preferred proxy for initial productivity is based on firm-level revenues per worker at the start of the analysis period (2000), obtained by linking the LBD to the BRFIRM_REV dataset.[39] I rank firms by their 2000 log revenues per worker within national 5-digit NAICS code by age group bins.[40] I use these ranks to assign firms to the deciles d seen in Equation (4.2).[41] Section 4.2.2 discusses the motivations, strengths, and weaknesses of this measure in more detail, but I start by describing the main results.

4.2.1 Main Results

Figure 6 plots coefficient estimates $\hat{\gamma}_d$ from Equation (4.2). The results lie in stark contrast to concerns of stunted creative destruction. Instead, they show a clear *increase* in establishment exit in the bottom quintile of the productivity distribution in response to immigrant inflows, with a particularly large increase in the lowest decile ($\hat{\gamma}_1 > \hat{\gamma}_2 > 0$). $\hat{\gamma}_d$ declines monotonically through $d = 6$, crossing zero between the second and third deciles, and remains flat for $d \in \{7, ..., 10\}$. Consistent

[38] See Appendix Section B.7.2.

[39] Establishment-level revenues are not available in my data.

[40] Firm age groups are 0-1, 2-4, 5-9, 10-19, and 20+. Firm-level NAICS codes are provided in the BRFIRM_REV dataset.

[41] Specifically, to accommodate (the few) 5-digit NAICS by age group pairs that contain fewer than 10 firms, a firm's decile is defined as Floor $\left(10 \times \frac{\text{Rank}_f - 1}{\text{Total}}\right) + 1$.

with the overall decrease in exit found in Sections 3.1.2 and 4.1, $\hat{\gamma}_d < 0$ for the majority of d, and $|\sum_d \mathbb{1}\{\hat{\gamma}_d < 0\}\hat{\gamma}_d| > \sum_d \mathbb{1}\{\hat{\gamma}_d > 0\}\hat{\gamma}_d$.

These results have novel implications for the heterogeneous impact of immigration in the U.S. and the source of this effect. Taken as a whole, these show that immigrant worker inflows help prevent exit for most establishments, but suggest that they may raise the productivity bar firms need to clear in order to continue operating establishments. In many models of firm heterogeneity, including the one presented in Section 5, raising this bar also increases aggregate productivity. I directly test for this implication in Section 4.3.

Figure 6 also represents a strong signal that immigrant workers are specifically tied to higher productivity firms—or at least, firms that are not in the lowest quintile. Section 5 also posits one potential reason this could take place: some firms can better utilize immigrant labor, but at a cost that makes them positively selected on productivity.[42] Regardless of the reason, these results imply that firm heterogeneity is critical to the understanding of immigration.

Finally, these results strongly comport with my assertion that the effects measured in this paper originate from increased immigrant labor supply, as opposed to increased immigrant consumer demand. An immigrant-induced increase in consumer demand would likely benefit, rather than cull, establishments from marginal, low productivity firms.

4.2.2 *Proxying for Productivity*

Foster et al. (2008) motivate both the use of the revenuer per worker productivity proxy and the adjustments I make to align it more closely with total factor, or "physical," productivity. First, Foster et al. (2008) find that revenue-per-input measures of productivity correlate strongly with measures of physical productivity in industries where these concepts can be separated cleanly. Second, their measures of revenue-based total factor productivity account for capital, materials, and energy inputs along with labor. I thus rank firms within detailed industries—where input requirements are more likely to be similar—to mute differences in non-labor input use from driving differences in my productivity measure. Ranking within detailed industry also removes cross-industry market-power differentials from driving revenue per worker differences. Finally, Foster et al. (2008) also find that divergence between revenue-based and physical measures of productivity—even within detailed industries—often occurs because of demand shocks that reflect market foothold. For example, older firms that produce the same product as younger firms may generate more demand because of non-quality-related factors like name recognition. This motivates ranking firms within age groups.

Nonetheless, a large literature starting with Klette and Griliches (1996) and including Foster et al. (2008) indicates that this measure still has clear limitations as a proxy for true physical productivity. Because I do not observe firm-level prices or non-labor inputs, I ultimately cannot fully

[42] Other possibilities include Becker-type discrimination among lower productivity firms, frictions that prevent lower productivity firms from directing search to immigrant workers they may want to hire, and greater adaptability of higher productivity firms.

eliminate the influence of these factors in my measure of productivity.[43] Ranking within industry-by-age bins is instead an attempt to improve upon the 0.75 correlation found between revenue-per-input and physical productivity measures in Foster et al. (2008). It should also be noted that in simplified models with competitive labor markets—including the one presented below—revenues per worker are independent of total factor productivity. These models, instead, have a direct link between firm size (both employment and revenues) and firm total factor productivity. I therefore test the robustness of my results to the use of size-based productivity proxies in Section 4.2.3.

The key assumption can be stated as: ranking within 5-digit NAICS code and age groups and controlling for time-invariant establishment unobservables (α_e) removes enough influence from idiosyncratic demand and non-labor input use across firms such that remaining variation in revenues per worker primarily reflects differences in physical productivity. The plausibility of this assumption is strengthened by robustness of results to using alternate measures that also correlate with physical productivity.

As additional motivation, I also note the prima facie interest understanding how immigrant worker inflows affect firms along the revenue per worker, revenue, employment, and payroll per worker distributions. At best, these also measure the impact of immigrant worker inflows on the total factor productivity distribution. At worst, they measure important features of the firms and establishments that operate in local economies. For example, from the perspective of a worker, whether high- or low-paying firms remain in the market in response to immigrant worker inflows is in it of itself a critical question.

4.2.3 Robustness

To increase power and provide several robustness checks around the results presented in Figure 6, I take advantage of its smooth, convex shape and estimate models using the following specification:

$$\begin{aligned}\mathbb{1}[\text{Not Operating}]_{et} =& \eta_0 \left(I_{gkt}\right) + \eta_1 \left(I_{gkt} \times [\text{Prod. Pctl.}]_e\right) + \eta_2 \left(I_{gkt} \times [\text{Prod. Pctl.}]_e^2\right) \\ &+ \Gamma X_{gkt} + \alpha_e + \alpha_{gt} + \alpha_{r(g),kt} + \varepsilon_{et}\end{aligned} \tag{4.3}$$

where Prod Pctl. is simply the continuous rank percentile of a given establishment in a given productivity distribution.[44]

$[\text{Prod Pctl.}]_e$ is measured in several ways, differing along three dimensions: 1) the unit for which the productivity measure is defined—either firm (as above) or establishment; 2) the bin within which that unit is ranked—either national NAICS 5-digit industry by age group (as above) or simply within local industry gk; and 3) the productivity measure itself—revenues per worker (as above), revenues, employment or payroll per worker. When productivity measures are defined at the firm level, all establishments in that firm are assigned the same $[\text{Prod Pctl.}]_e = [\text{Prod Pctl.}]_{f(e)}$. As a starting point, I estimate Equation (4.3) using firm-level revenues per worker ranked within

[43] Indeed, Foster et al. (2008) find that these factors are important even within very detailed industries

[44] Specifically, Pctl. is defined as $100 \times \frac{\text{Rank}_f - 1}{\text{Total}}$.

national NAICS 5-digit industry by age group bins, as in Section 4.2.1. Figure A3 in the Appendix demonstrates that the estimates from this model are consistent with the non-parametric results from Section 4.2.1.

Table 4 displays the results of these robustness exercises, estimated using the vector of instruments $\left(z_{gkt}^{\text{Emigrants}}, z_{gkt}^{\text{Emigrants}} \times [\text{Prod. Pctl.}]_e, z_{gkt}^{\text{Emigrants}} \times [\text{Prod. Pctl.}]_e^2\right)$ to estimate $\hat{\eta}_0$, $\hat{\eta}_1$, and $\hat{\eta}_2$. For ease of viewing coefficient estimates, productivity percentiles have been divided by 100. The "Implied Crossing Point Pctl." row displays the percentile at which the effect of immigrant inflows on exit turns from positive to negative. Across productivity measures and control sets, all columns show the same pattern: culling at the low end of the productivity distribution, and a convex drop that crosses over to exit prevention somewhere between the 27th and 38th percentile. In conjunction, $\hat{\eta}_0$ and the "Implied Crossing Point Pctl." row reveal differences in the degree of low productivity culling across specifications.

Several specific results are worth noting, starting with a comparison between Columns 1 and 2. Column 1 displays the results that generate the fit in Figure A3 and stand in for the "Main Results" in Table 4. Column 2 demonstrates that these main results are robust to the inclusion of an endogenous control for native worker presence, along with its interaction with the linear and squared productivity terms. Increases in exposure to immigrant workers tend to have a much larger impact on establishment exit decisions than increases in exposure to native worker inflows. Furthermore, native worker inflows appear to have the *opposite* effect across the productivity distribution, with small preservation effects on establishments from the lowest productivity firms. This bolsters the case for immigrant-specific ties to high productivity firms. In addition, native worker displacement does not appear to drive the results in Column 1, and the stability of the 1st Stage F statistic across the two columns indicates that the instrument is appropriately acting through immigrant workers rather than general workforce growth.

Columns 3 through 7 indicate robustness to alternate productivity measures. Columns 3 through 5 continue to define productivity at the firm-level within industry-age bins, and show that the pattern remains in place when we use size-based productivity measures (revenues or employment) and when we use measures that are available for the full set of establishments (employment and payroll per worker). Columns 6 and 7 simply rank establishments at the level of the exposure variable—the local industry—in terms of employment and payroll per worker (average annual earnings). They once again display the same broad pattern, although Column 6 shows a more linear functional form and higher threshold at which culling still occurs. A worker in a receiving local industry can expect inflows of immigrant workers to cull the smallest and lowest paying (on average) establishments, while sustaining the largest and highest paying establishments.

Appendix Table A6 presents heterogeneity analyses across industry groups (k) and regions ($r(g)$) using firm-level revenues per worker as the productivity measure (for comparison with Column 1 of Table 4).

4.3 Immigrant Worker Inflows and the Establishment Productivity Distribution

Figure 6 and Table 4 introduce stratified, establishment exit as a novel mechanism through which immigrant workers affect local U.S. industries. These responses also introduce a new channel through which immigrant worker inflows can directly affect productivity in a local industry: by changing the composition of firms that operate in it. Next, I directly estimate how immigrant worker inflows impact the count of establishments operating across the productivity distribution in a local industry.

Specifically, I estimate a modified version of Equation (2.1) that takes one long difference over the time period 2000—2015

$$\Delta y_{gk}^{d} = \beta^{d}\left(\Delta I_{gk}\right) + \Gamma X_{gk} + \alpha_{g} + \alpha_{r(g),k} + \varepsilon_{gk} \tag{4.4}$$

where Δ now represents a difference between 2000 and 2015. Outcomes once again come from a decomposition of the DHS growth rate in establishment count, this time by productivity:

$$\Delta y_{gk}^{d} = \frac{\text{Estabs}_{gk,t=2015}^{d} - \text{Estabs}_{gk,t=2000}^{d}}{\left(\frac{\text{Estabs}_{gk,t=2015} + \text{Estabs}_{gk,t=2000}}{2}\right)}$$

where d indexes a productivity decile. To construct Δy_{gk}^{d}, I start by ranking firms within 5-digit NAICS code by age group bins in 2000, as above. Using these rankings, I obtain cutoffs that define each 5-digit NAICS code by age group bin's deciles. Δy_{gk}^{d} then counts the change in the number of operating establishments whose parent firm's real revenues per worker fall within the 2000-based decile d for their 5-digit NAICS code and age group. Estimated coefficients include the impact of immigrant worker inflows on establishment exit—as studied in detail in Section 4.2.1—but also establishment entry and continuing establishment movement across productivity bins. Furthermore, $\sum_{d=1}^{10} \beta^{d}$ gives the overall effect of immigrant worker inflows on the DHS growth rate in establishment presence over the time period 2000–2015. The instrumental variable, $\Delta z_{gk}^{\text{Emigrants}}$, takes the same form as in Section 3.1.1, but differencing between 2000 and 2015.

The results presented in Figure 7 provide a simple and powerful summary of the extensive margin effects immigrant workers have on local industry productivity. There are small increases in the number of operating establishments from firms with revenues per worker in the lowest seven deciles of the productivity distribution. Meanwhile, there are large and growing effects in the top three deciles, with the top decile accounting for 26 percent of the overall increase in the DHS growth rate and the top three deciles accounting for 63 percent. This is a substantive change in the distribution of firms that operate in a local industry, heavily tilted towards higher productivity firms. The overall IV-estimated effect, adding up estimated coefficients, is a 0.54 percent increase in the DHS growth rate in establishment count for a 1 percent shock to relative supply induced by immigrant worker inflows.

5 A Synthesizing Framework

The previous two sections have delivered key empirical insights into how immigrant workers affect local economies in partial equilibrium (i.e., comparing across sectors within the same commuting zone). This section relies on theory to frame how accounting for the kind of firm heterogeneity implied by the partial equilibrium results ultimately impacts general equilibrium analyses of immigration. Appendix Section C provides an expanded version of the model and uses it to quantify the importance of differences across firms to the welfare impact of immigration.

5.1 Motivation

By definition, models featuring perfect competition are not compatible with firm heterogeneity in the final goods markets, and models featuring both perfect competition and perfect substitutability in the market for foreign-born and domestic labor are not compatible with immigrant-specific effects on firm outcomes. Melding the Melitz (2003) model with imperfect substitutability in the labor market across immigrant and native workers allows immigrant employees to generate distinct effects on firm outcomes relative to native employees and potentially change the productivity distribution of firms.

However, this combination alone does not capture the specific linkages between higher productivity firms and immigrant employees implied by the results in Section 4. In order to capture productivity-related reallocative responses to immigration, I thus add an additional heterogeneity to my theoretical framework. In the spirit of Bustos (2011), I introduce two technology types—one of which is more suited to utilize immigrant labor. Firms must pay an additional fixed operating cost to access this technology. These costs can include (but are hardly limited to) hiring translators and liaisons to be able to enter into immigrant job search networks and direct search for immigrants,[45] hiring lawyers to work on visa issues, and paying enforcement costs (in expectation) when hiring undocumented immigrants.

In the presence of such costs, firms that utilize the immigrant-intensive technology will be positively selected on productivity relative to their counterparts due to a spreading effect. Immigrant-intensive firms save more on labor costs when immigrants enter the labor market, pass through their savings to consumers, and thus gain market share by competing it away from firms who do not have special ties to immigrant workers. Their counterparts are forced to exit, a culling of the lowest-productivity firms in the market as a whole. In sum, this model shows how firm entry and exit can drive immigrant-induced endogenous technological change by changing the composition of firms in the economy.

[45] See e.g., this Center for American Progress report about Tyson Fresh Meats and its willingness to hire translators, liaisons, and chaplains in order to utilize immigrant labor.

5.2 Setup

Individuals are consumer-employees of type $i \in \{I, N\}$, with I representing the foreign-born and N representing the native-born. The mass of each labor type in the economy is fixed and employees supply their labor inelastically—the primary comparative static will increase immigrant mass I.

5.2.1 *Firms*

The market structure is monopolistic competition, and each firm indexes a product.[46] An endogenous mass of potential entrepreneurs pay an entry cost to take a draw of total factor productivity z from a known Pareto distribution with shape parameter ϕ and minimum value m. This productivity draw endogenously determines whether they start a firm, and if so, with which technology type $j \in \{0, 1\}$. Firm production functions are given by

$$Q_j(z) = zq_j(z)$$
$$q_j(z) = \left[a_j\,(I(z))^{\frac{\sigma-1}{\sigma}} + (N(z))^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$$

where $Q_j(z)$ is total production of a firm whose owner draws productivity z and production with technology type j. $q_j(z)$ is a CES aggregator of immigrant and native labor employed by the firm—the only two factors or production. σ is the elasticity of substitution between immigrant and native workers. The key difference across firms producing with technology j is the parameter a_j. $j = 1$ firms depend more on and better-utilize immigrant labor:

$$a_1 > a_0$$

The cost function is given by

$$\left(\frac{c_j}{z}\right) Q_j(z) + c_j \kappa_j^f$$

where

$$c_j \equiv \left[a_j^{\sigma}\,(w_I)^{1-\sigma} + (w_N)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}$$

and w_i represents the wage for a worker of nativity i. κ_j^f is a fixed operating cost that varies by technology type, as motivated above. Entrepreneurs must decide whether or not to pay a proportional cost every period, $\tau > 1$, such that $\kappa_1^f = \tau\kappa_o^f$ in order to access the immigrant-specific production boost represented by $a_1 > a_0$. I normalize $w_N = 1$.

5.2.2 *Consumer Preferences*

Consumer preferences have CES form with elasticity of substitution μ such that $\phi > \mu - 1 > 0$: :

[46] There is no distinction between a firm and an establishment in the model. Each firm/establishment represents an additional variety.

$$\mathcal{U} = \left[F^{\frac{\eta-1}{\mu}} \int_0^F Q(f)^{\frac{\mu-1}{\mu}} df \right]^{\frac{\mu}{\mu-1}}$$

F is the mass of firms in the economy and $Q(f)$ is the amount demanded by consumers from firm f. When $\eta = 1$, consumers have a taste for variety, which generates external scale effects. When $\eta = 0$, we shut down this channel from market size to welfare and focus on the firm productivity distribution (see, e.g., Egger and Kreickemeier, 2009).

This results in the following demand curves for each firm, which are downward sloping due to product differentiation and substitutability across goods:

$$Q(f) = YF^{\eta-1}P^{\mu-1}p(f)^{-\mu} \tag{5.1}$$

where $p(f)$ is the price charged by firm f , Y is total consumer spending, and the price index P is given by $P^{1-\mu} \equiv F^{\eta-1} \int_0^F p(f)^{1-\mu} df$.

5.2.3 *Prices*

This setup leads to a familiar pricing rule in models with CES preferences:

$$p_j(z) = \left(\frac{\mu}{\mu - 1}\right)\left(\frac{c_j}{z}\right) \tag{5.2}$$

That is, the firm still charges a constant markup over its marginal cost, but the firm's marginal cost reflects the two different types of labor it aggregates. Firms compete through prices, and so declines in c_j help firms gain market share.

5.3 Equilibrium Conditions

Let $\pi_j(z)$ index profits for the firm with productivity z producing with technology type j. Following the logic of Bustos (2011), there exists a cutoff, z_1^* at which marginal producers are indifferent between the high immigrant-use and low-immigrant use technologies:

$$\pi_0(z_1^*) \equiv \pi_1(z_1^*) \tag{5.3}$$

Firms with productivity draws below this cutoff produce with technology $j = 0$, while firms above this cutoff produce with technology $j = 1$. This is due to scale. Firms with higher productivity draws produce more output and require a larger workforce. These larger firms find it profitable to pay a larger fixed cost in order to reduce their variable costs because it is spread across more employees. Firms with lower productivity draws do not produce at a scale that justifies paying this larger fixed cost.

Entrepreneurs only stay in the market if they are profitable. This defines the usual cutoff productivity for type 0 firms:

$$\pi_0(z_0^*) \equiv 0 \tag{5.4}$$

So, entrepreneurs with productivities below z_0^* exit the market, entrepreneurs with productivities in $[z_0^*, z_1^*]$ produce with technology 0, and entrepreneurs with productivities above z_1^* produce with technology 1. Entrepreneurs do not know their z prior to entry, and must pay an entry cost to discover it.

Free entry yields

$$\mathbb{E}[\pi(z)] = \mathbb{E}[\pi(z)|z > z_0^*]\mathbb{P}[z > z_0^*] = c_0\kappa^e \tag{5.5}$$

where κ^e is the sunk (entry) cost potential entrepreneurs pay to take productivity draws, denominated in units of output. When profits are high enough, entrepreneurs enter until they no longer expect to recover their entry costs.[47]

Finally, the price level, P is given by

$$P^{1-\mu} \equiv n_e \left[\int_{z_0^*}^{z_1^*} p_0(z)^{1-\mu} g(z)dz + \int_{z_1^*}^{\infty} p_1(z)^{1-\mu} g(z)dz\right] \tag{5.6}$$

5.4 Key Insights

5.4.1 The Profitability Cutoff

The key item of interest is z_0^*. Solving Equations (5.3) through (5.6) yields

$$z_0^* = m\left[\left(\frac{\kappa_0^f}{\kappa^e}\right)\left(\frac{\mu-1}{\phi-(\mu-1)}\right)\theta\right]^{\frac{1}{\phi}} \tag{5.7}$$

$$\theta \equiv 1 + \left(\frac{z_1^*}{z_0^*}\right)^{-\phi}\left(\frac{c_1\tau - c_0}{c_0}\right) \tag{5.8}$$

That is, the profitability cutoff z_0^* is constant except for θ, an endogenous variable that sets this model apart from more standard models of firm heterogeneity (e.g., Melitz, 2003). It introduces the notion that entry and exit decisions for marginal, $j = 0$ firms depend on inframarginal, $j = 1$ firms, through their ability to steal away market share when their costs go down.

Appendix Section C shows that under reasonable calibrations, θ generates the following dynamics: when immigrant workers enter the labor market, relative immigrant wages w_I fall, and firms producing with type $j = 1$ technology see their labor costs fall by more than firms producing with type $j = 0$ technology. $j = 1$ firms pass these labor cost savings on to consumers through prices and are thus able to compete away some of the market from $j = 0$ firms. Reduced demand raises the productivity bar, z_0^*, that $j = 0$ firms must cross in order to be profitable. The distinction between partial and general equilibrium notwithstanding, evidence from Section 4.2 is consistent with these dynamics, with immigrant worker inflows culling establishments from firms

[47] Note that the assumption that entry costs scale with c_0 (instead of c_1 or a combination) is mostly made for analytical convenience. However, a simple, plausible justification is that producers do not invest in the costs to access immigrant labor until after entry activities have been completed and they find out they have a draw of z above z_1^*. Thus, the entry activities are paid for using type 0 technology.

at the lower tail of the productivity distribution.

5.4.2 *Why does the Profitability Cutoff Matter? The Immigrant Surplus*

Solving Equations (5.3) through (5.6) also yields

$$P^{1-\mu} = [\text{Const.}](c_0)^{1-\mu} F^{\eta} (z_0^*)^{\phi+(\mu-1)}$$

Given that native wages are the numeraire, we have

$$-\frac{d\log(w_N/P)}{dI} = \underbrace{-\frac{d\log(c_0)}{dI}}_{\text{Analogous to rep. firm model}} + \underbrace{\left(\frac{\eta}{\mu-1}\right)\frac{d\log(F)}{dI}}_{\text{Increased variety}} + \underbrace{\left(\frac{\phi}{\mu-1}+1\right)\frac{d\log(z_0^*)}{dI}}_{\text{Culling of marginal firms}} \tag{5.9}$$

Equation (5.9) clarifies the value added of this modeling framework by decomposing the "immigrant surplus"—the effect of an immigrant inflow on real native incomes. First, as shown in Appendix Section D.1, a canonical, representative firm model of production in which all firms have access to type 0 technology would yield an immigration surplus equal to the first term. These are second-order effects that are the result of relative wage changes across natives and immigrants (Dustmann et al., 2012). The additional terms are driven by extensive margin changes and would be missed by this canonical framework. When $\eta = 1$, natives benefit through increased variety, a channel explored in di Giovanni et al. (2014) and Hong and McLaren (2015).[48]

The third and final term in Equation (5.9) is novel to the literature and hones in on the behavior of firms at the exit margin. It changes the effect of immigrant workers on native welfare even when we ignore variety ($\eta = 0$). An immigrant-induced rise in z_0^* imparts a more than one-for-one increase in real native incomes. As the composition of firms is selected on higher levels of productivity, consumers see lower prices through the pass-through implied by Equation (5.2). This term implies that examining the response of lower-productivity firms can be a critical informant of a first-order effect of immigration on an economy. Results in Section 4.2 imply that z_o^* is rising in partial equilibrium. Simulation exercises in Appendix Section C show that this channel is quantitatively important in general equilibrium as well.

This decomposition, taken together with the empirical results presented in this paper, shows how accounting for heterogeneous firm responses can lead to estimates of immigrant-generated welfare that are substantially larger than those that come from canonical models of labor demand. By allowing for dispersion of productivity across firms, models with the Melitz (2003) structure open a channel to welfare through the productivity distribution. In particular, a rising productivity cutoff increases welfare by lowering prices—a *first-order* welfare gain due to changing firm composition.[49] Welfare gains that arise from changes to wages and firm input costs, by contrast,

[48] Note that these variety gains would not actually show up in measured native "incomes" because price indices generally do not account for increasing variety.

[49] Yet another first-order welfare gain—not modeled here—could arise if competition among firms raised the elasticity of substitution across products (see, e.g., Blanchard and Giavazzi, 2003).

are small because the labor market is characterized by perfect competition.

5.5 Further Evidence for Mechanisms

The model described above implies that firms that employ more immigrants in production 1) tend to be positively selected on productivity and 2) see larger labor cost reductions in response to immigrant worker inflows. Here, I present evidence consistent with the implications and mechanisms of the model.

First, related literature from Europe provides evidence that firms that employ immigrant workers tend to be more productive. Mitaritonna et al. (2017) find that manufacturing firms in France that employ immigrant workers are more than 11 percent more productive than those that employ zero immigrants. Brinatti and Morales (2021) find that larger firms in Germany have larger immigrant shares of their workforce.[50] Next, I provide additional evidence by further characterizing stratified, establishment-level responses to immigrant worker inflows in 2000-2015.

5.5.1 Which Establishments See Labor Cost Reductions in Response to Immigration?

The productivity-enhancing dynamics in the model presented above hinge on labor cost savings that accrue to higher-productivity firms when there is an influx of immigrant workers, due to the importance of immigrant workers in their production technology. To test whether these labor cost savings take place, I employ the following specification on the panel of establishments that were operating as of 2000 and can be linked to revenue information, from Section 4.2. I modify Equation (4.2) to study average earnings at an establishment, conditional on remaining in operation:[51]

$$\log(\text{Payroll p.w.})_{et} = \sum_{d=1}^{10} \rho_d \left(I_{gkt} \times \mathbb{1}\{f(e) \in d\} \right) + \Gamma X_{gkt} + \alpha_e + \alpha_{gt} + \alpha_{r(g),kt} + \varepsilon_{et} \tag{5.10}$$

Figure 8 plots the resulting coefficient estimates $\hat{\rho}_d$, using $z_{gkt}^{\text{Emigrants}}$ as the instrumental variable for I_{gkt}. A stark pattern emerges in which establishments from firms in the top three deciles of the productivity distribution—particularly the top decile—see significant reductions in labor costs, while those throughout the rest of the productivity distribution generally see labor cost increases. These results imply that labor cost savings are central to the the extensive margin responses detected in Section 4 and validate labor costs as a key building block in a model of immigration and firm heterogeneity.

[50] Firm size is a one-for-one correlate with productivity in the model presented above.

[51] Note that payroll includes both native and immigrant earnings, so ρ_d^{Earn} contains changes to payroll due to changes in wages due to changes in firm- or market-level productivity of all workers, along with changes in group-specific wages weighted by workforce nativity-group composition.

5.5.2 *Stratified Exit Responses by Firm Ownership Nativity*

In the model presented above, low productivity firms exit the market because they do not pay the fixed costs required to employ a large immigrant share in their workforce. However, immigrant firm-owners likely face lower fixed costs to setting up immigrant-hiring apparatuses through knowledge of co-ethnic networks and the immigration system (Kerr and Kerr, 2021). In this case, lower-productivity, immigrant-owned firms should be somewhat buffered from the immigrant-inflow-induced market forces that cull low productivity firms from the market. Meanwhile, higher productivity firms—regardless of ownership nativity—should exhibit behavior consistent with having a larger immigrant share in their workforce.

In order to compare extensive margin, establishment-level responses to immigrant worker inflows stratified both by productivity and by ownership nativity, I link the 2007 Survey of Business Owners (SBO) to the LBD.[52] The 2007 SBO covers a representative sample of operating firms in 2007 and included questions about firm owner nativity for the three individuals with highest ownership stakes of each firm.[53] Letting a firm be immigrant owned if at least 50 percent of its ownership stake was owned by those born outside of the U.S., I take the set of operating SBO establishments in 2007 and follow them for $t \in \{2007, 2012, 2017\}$ using the following specification:

$$\begin{aligned}
\mathbb{1}\{\text{Not Operating}\}_{et} = &\sum_{d=1}^{10} \gamma_d^I \left(I_{gkt} \times \mathbb{1}\{\text{Immigrant-Owned}_{f(e)} = 1\} \times \mathbb{1}\left\{ f(e) \in d \right\} \right) \\
&+ \sum_{d=1}^{10} \gamma_d^N \left(I_{gkt} \times \mathbb{1}\{\text{Immigrant-Owned}_{f(e)} = 0\} \times \mathbb{1}\left\{ f(e) \in d \right\} \right) \\
&+ \Gamma X_{gkt} + \alpha_e + \alpha_{gt} + \alpha_{r(g),k,t} + \varepsilon_{et}
\end{aligned} \tag{5.11}$$

where deciles d are defined using ranks of sales per worker within 5-digit NAICS code by age group bins.[54] Due to the timing of the UNPD emigration data, I use $z_{gk,t-2}^{\text{Emigrants}}$ as the instrumental variable for this specification. Appendix Table A7 shows that this sample of firms and alternate time period generates the same overall pattern found in Section 4.2 of low productivity culling and medium and high productivity preservation.

Estimated coefficients from (5.11) are presented in Figure 9. Although the lowest-productivity firms of each nativity group are culled from the market, there is a clear distinction in the responses of immigrant- and native-owned firms in the first productivity decile, with the latter more than 60 percent more likely to exit the market in response to immigrant worker inflows. Primarily because of this distinction, the overall pattern of low productivity culling is more muted for immigrant-owned firms, consistent with a fixed-cost mechanism that is stratified by nativity. This strongly implies that there are linkages between immigrant entrepreneurs and immigrant employees.[55]

[52] For this analysis, I used the Revised LBD, a new Census product that includes years needed for this analysis. The original LBD, used in the rest of this paper, stops in 2016.

[53] Publicly-traded firms are excluded from this analysis.

[54] Total sales are included for all SBO firms. This precludes the need for the BRFIRM_REV dataset in this analysis.

[55] Recall that α_{gt} limits the influence of immigrant consumer demand.

These results also indicate that the increased dynamism found in Section 4 is not without distributional consequences. Just as the literature on how immigration affects employees has focused on natives at risk of being substituted for in production, it appears that there is a subset of natives who own low productivity firms that are most at risk of having to shut down their business in response to increased immigrant presence. In the face of immigrant inflows, immigrant entrepreneurs' ties to immigrant workers translate to a competitive advantage over their native counterparts.

6 Conclusion

This paper documents several empirical relationships that reveal the critical role establishment entry and establishment exit play in both absorbing immigrants into and mediating the effects immigrant workers have on local industries. Both entry and exit account for an economically significant, positive relationship between immigrants and establishment presence over three decades in which immigration was a defining feature of demographic change in the U.S. Exit prevention plays an overwhelming role in the net job creation that absorbs immigrant worker inflows into local industries. Contrary to what we would expect if these effects were driven solely by consumer demand or uniformly lower labor costs across firms, immigrant inflows cull establishments from lower productivity firms. Consequently, immigrant worker inflows increase the number of operating establishments in the right tail of the productivity distribution. Ties between immigrant workers and higher productivity firms imply a substantially larger "immigration surplus" than we would estimate with standard, representative firm models of the product market. In fact, responses by firm entry and exit—and subsequent changes to firm composition—represent a plausible channel through which the immigrant surplus can be first-order.

Future work will delve into and contextualize the link between immigrant workers, business entry and exit, and higher-productivity firms. Employer-employee linked data, for example, can test the model's assertion that more productive firms within U.S. industries tend to both hire more immigrant workers and benefit more from immigrant inflows. Additionally, cross-geographic comparisons that stratify immigrant absorption outcomes—including wage changes, native displacement, and productivity—by variables that reflect the ease of starting and shutting down a business can validate this paper's suggestion that flexibility on the extensive margin is a key determinant of the relative success in U.S. immigrant labor market assimilation.

References

Adao, Rodrigo, Michal Kolesar, and Eduardo Morales, "Shift-Share Designs: Theory and Inference," *The Quarterly Journal of Economics*, 08 2019, *134* (4), 1949–2010.

Altindag, Onur, Ozan Bakis, and Sandra V. Rozo, "Blessing or burden? Impacts of refugees on businesses and the informal economy," *Journal of Development Economics*, 2020, *146*, 102490.

Autor, David H. and David Dorn, "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, August 2013, *103* (5), 1553–97.

—, —, and Gordon H. Hanson, "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, October 2013, *103* (6), 2121–68.

Ayromloo, Shalise, Benjamin Feigenberg, and Darren Lubotsky, "States Taking the Reins? Employment Verification Requirements and Local Labor Market Outcomes," NBER Working Papers 26676, National Bureau of Economic Research, Inc January 2020.

Azoulay, Pierre, Benjamin Jones, J. Daniel Kim, and Javier Miranda, "Immigration and Entrepreneurship in the United States," Working Paper 27778, National Bureau of Economic Research September 2020.

Bartelsman, Eric J. and Mark Doms, "Understanding Productivity: Lessons from Longitudinal Microdata," *Journal of Economic Literature*, September 2000, *38* (3), 569–594.

Bartik, Timothy, *Who Benefits from State and Local Economic Development Policies?*, Upjohn Institute, 1991.

Beerli, Andreas, Jan Ruffner, Michael Siegenthaler, and Giovanni Peri, "The Abolition of Immigration Restrictions and the Performance of Firms and Workers: Evidence from Switzerland," *American Economic Review*, 2021.

Blanchard, Olivier and Francesco Giavazzi, "Macroeconomic Effects of Regulation and Deregulation in Goods and Labor Markets*," *The Quarterly Journal of Economics*, 08 2003, *118* (3), 879–907.

Borjas, George J., "The economic analysis of immigration," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Vol. 3 of *Handbook of Labor Economics*, Elsevier, May 1999, chapter 28, pp. 1697–1760.

—, *Immigration Economics*, Harvard University Press, 2014.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel, "Quasi-Experimental Shift-Share Research Designs," Working Paper 24997, National Bureau of Economic Research September 2020.

Brinatti, Agostina and Nicolas Morales, "Firm Heterogeneity and the Impact of Immigration: Evidence from German Establishments," 2021. Working Paper.

Brücker, Herbert, Stella Capuano, and Abdeslam Marfouk, "Education, gender and international migration: insights from a panel-dataset 1980-2010," *Methodology Report*, 2013.

Burchardi, Konrad B, Thomas Chaney, Tarek Alexander Hassan, Lisa Tarquinio, and Stephen J Terry, "Immigration, Innovation, and Growth," Working Paper 27075, National Bureau of Economic Research May 2020.

Burstein, Ariel, Gordon Hanson, Lin Tian, and Jonathan Vogel, "Tradability and the Labor-Market Impact of Immigration: Theory and Evidence From the United States," *Econometrica*, 2020, *88* (3), 1071–1112.

Bustos, Paula, "Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms," *American Economic Review*, February 2011, *101* (1), 304–40.

Cadena, Brian C. and Brian K. Kovak, "Immigrants Equilibrate Local Labor Markets: Evidence from the Great Recession," *American Economic Journal: Applied Economics*, 2016.

Christopoulou, Rebekka and Philip Vermeulen, "Markups in the Euro area and the US over the period 1981–2004: a comparison of 50 sectors," *Empirical Economics*, Feb 2012, *42* (1), 53–77.

Clemens, Michael A., Ethan G. Lewis, and Hannah M. Postel, "Immigration Restrictions as Active Labor Market Policy: Evidence from the Mexican Bracero Exclusion," *American Economic Review*, June 2018, *108* (6), 1468–87.

Cortes, Patricia, "The Effect of Low-Skilled Immigration on U.S. Prices: Evidence from CPI Data," *Journal of Political Economy*, 2008, *116* (3), 381–422.

D'Amuri, Francesco, Gianmarco Ottaviano, and Giovanni Peri, "The labor market impact of immigration in Western Germany in the 1990s," *European Economic Review*, 2010, *54* (4), 550–570.

di Giovanni, Julian, Andrei A. Levchenko, and Francesc Ortega, "A Global View of Cross-Border Migration," *Journal of the European Economic Association*, 2014, *13* (1), 168–202.

Doms, Mark, Ethan Lewis, and Alicia Robb, "Local labor force education, new business characteristics, and firm performance," *Journal of Urban Economics*, 2010, *67* (1), 61 – 77. Special Issue: Cities and Entrepreneurship.

Dustmann, Christian and Albrecht Glitz, "How Do Industries and Firms Respond to Changes in Local Labor Supply?," *Journal of Labor Economics*, 2015, *33* (3), 711–750.

, Tommaso Frattini, and Ian P. Preston, "The Effect of Immigration along the Distribution of Wages," *The Review of Economic Studies*, 04 2012, *80* (1), 145–173.

Eckert, Fabian, Teresa C Fort, Peter K Schott, and Natalie J Yang, "Imputing Missing Values in the US Census Bureau's County Business Patterns," Working Paper 26632, National Bureau of Economic Research January 2020.

Egger, Hartmut and Udo Kreickemeier, "Firm Heterogeneity and the Labor Market Effects of Trade Liberalization," *International Economic Review*, 2009, *50* (1), 187–216.

Fairlie, Robert W. and Magnus Lofstrom, "Chapter 17 - Immigration and Entrepreneurship," in Barry R. Chiswick and Paul W. Miller, eds., *Handbook of the Economics of International Migration*, Vol. 1 of *Handbook of the Economics of International Migration*, North-Holland, 2015, pp. 877 – 911.

Fort, Teresa C. and Shawn D. Klimek, "The Effects of Industry Classification Changes on US Employment Composition," Working Papers 18-28, Center for Economic Studies, U.S. Census Bureau June 2018.

Foster, Lucia, John Haltiwanger, and Chad Syverson, "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?," *American Economic Review*, March 2008, *98* (1), 394–425.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, "Bartik Instruments: What, When, Why, and How," *American Economic Review*, August 2020, *110* (8), 2586–2624.

Gonzalez, Libertad and Francesc Ortega, "How do very open economies adjust to large immigration flows? Evidence from Spanish regions," *Labour Economics*, January 2011, *18* (1), 57–70.

Haltiwanger, John, Ron Jarmin, Robert Kulick, Javier Miranda, Veronika Penciakova, and Cristina Tello-Trillo, "Firm-level Revenue Dataset," CES Technical Notes Series 19-02, Center for Economic Studies, U.S. Census Bureau August 2019.

Hong, Gihoon and John McLaren, "Are Immigrants a Shot in the Arm for the Local Economy?," Working Paper 21123, National Bureau of Economic Research April 2015.

Hopenhayn, Hugo, Julian Neira, and Rish Singhania, "From Population Growth to Firm Demographics: Implications for Concentration, Entrepreneurship and the Labor Share," Working Paper 25382, National Bureau of Economic Research December 2018.

Jaeger, David A, Joakim Ruist, and Jan Stuhler, "Shift-Share Instruments and the Impact of Immigration," Working Paper 24285, National Bureau of Economic Research February 2018.

Karahan, Fatih, Benjamin Pugsley, and Aysegul Sahin, "Demographic Origins of the Startup Deficit," Working Paper 25874, National Bureau of Economic Research May 2019.

Kerr, Sari Pekkala and William R. Kerr, "Immigrant Entrepreneurship," Working Paper 22385, National Bureau of Economic Research July 2016.

and , "Immigrant Entrepreneurship in America: Evidence from the Survey of Business Owners 2007 & 2012," Working Paper 24494, National Bureau of Economic Research April 2018.

— **and** — , "Whose Job Is It Anyway? Co-Ethnic Hiring in New U.S. Ventures," Working Paper 28509, National Bureau of Economic Research February 2021.

Klette, Tor Jakob and Zvi Griliches, "The inconsistency of common scale estimators when output prices are unobserved and endogenous," *Journal of Applied Econometrics*, 1996, *11* (4), 343–361.

Lewis, Ethan G., "Immigration, skill mix, and the choice of technique," Working Papers 05-8, Federal Reserve Bank of Philadelphia 2005.

— , "Immigrant-Native Substitutability: The Role of Language Ability," Working Paper 17609, National Bureau of Economic Research November 2011.

— , "Immigration and Production Technology," Working Paper 18310, National Bureau of Economic Research August 2012.

Llull, Joan, "The effect of immigration on wages: exploiting exogenous variation at the national level," *Journal of Human Resources*, 2017, pp. 0315–7032R2.

Melitz, Marc J., "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 2003, *71* (6), 1695–1725.

Mitaritonna, Cristina, Gianluca Orefice, and Giovanni Peri, "Immigrants and firms' outcomes: Evidence from France," *European Economic Review*, 2017, *96*, 62 – 82.

Monras, Joan, "Immigration and Wage Dynamics: Evidence from the Mexican Peso Crisis," *Journal of Political Economy*, 2020, *128* (8), 3017–3089.

— , "Local Adjustment to Immigrant-Driven Labor Supply Shocks," *Journal of Human Capital*, 2021, *15* (1), 204–235.

Olney, William W., "Immigration and Firm Expansion," *Journal of Regional Science*, 2013, *53* (1), 142–157.

Orefice, Gianluca and Giovanni Peri, "Immigration and Worker-Firm Matching," Working Paper 26860, National Bureau of Economic Research March 2020.

Orrenius, Pia M., Madeline Zavodny, and Alexander Abraham, "The Effect of Immigration on Business Dynamics and Employment," IZA Discussion Papers 13014, Institute of Labor Economics (IZA) February 2020.

Ottaviano, Gianmarco I.P. and Giovanni Peri, "Immigration and National Wages: Clarifying the Theory and the Empirics," Working Paper 14188, National Bureau of Economic Research July 2008.

— **and** — , "Rethinking the Effect of Immigration on Wages," *Journal of the European Economic Association*, 2012, *10.*

Ottinger, Sebastian, "Immigrants, Industries, and Path Dependence," 2020. Mimeo, University of California, Los Angeles.

Peri, Giovanni, "The Effect Of Immigration On Productivity: Evidence From U.S. States," *The Review of Economics and Statistics*, 2012, *94* (1), 348–358.

— **and Chad Sparber**, "Task Specialization, Immigration, and Wages," *American Economic Journal: Applied Economics*, 2009, *1* (3).

Porter, Michael, "The Economic Performance of Regions," *Regional Studies*, 2003, *37* (6-7), 549–578.

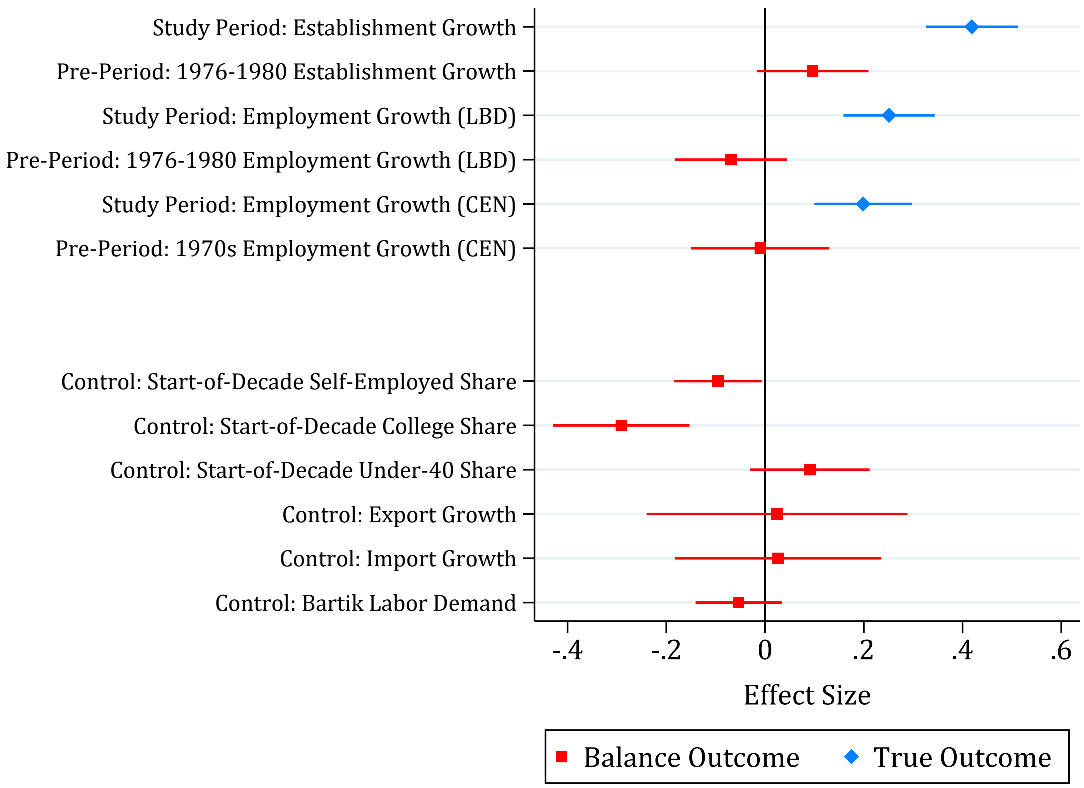
Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek, "IPUMS USA: Version 9.0 [dataset]," 2019.

Schoar, Antoinette, "The Divide between Subsistence and Transformational Entrepreneurship," in "Innovation Policy and the Economy, Volume 10," University of Chicago Press, February 2010, pp. 57–81.

Vespa, Jonathan and David M. Armstrong, "PS 25-1144: Demographic Turning Points for the United States: Population Projections for 2020 to 2060," Technical Report, U.S. Census Bureau 2018.

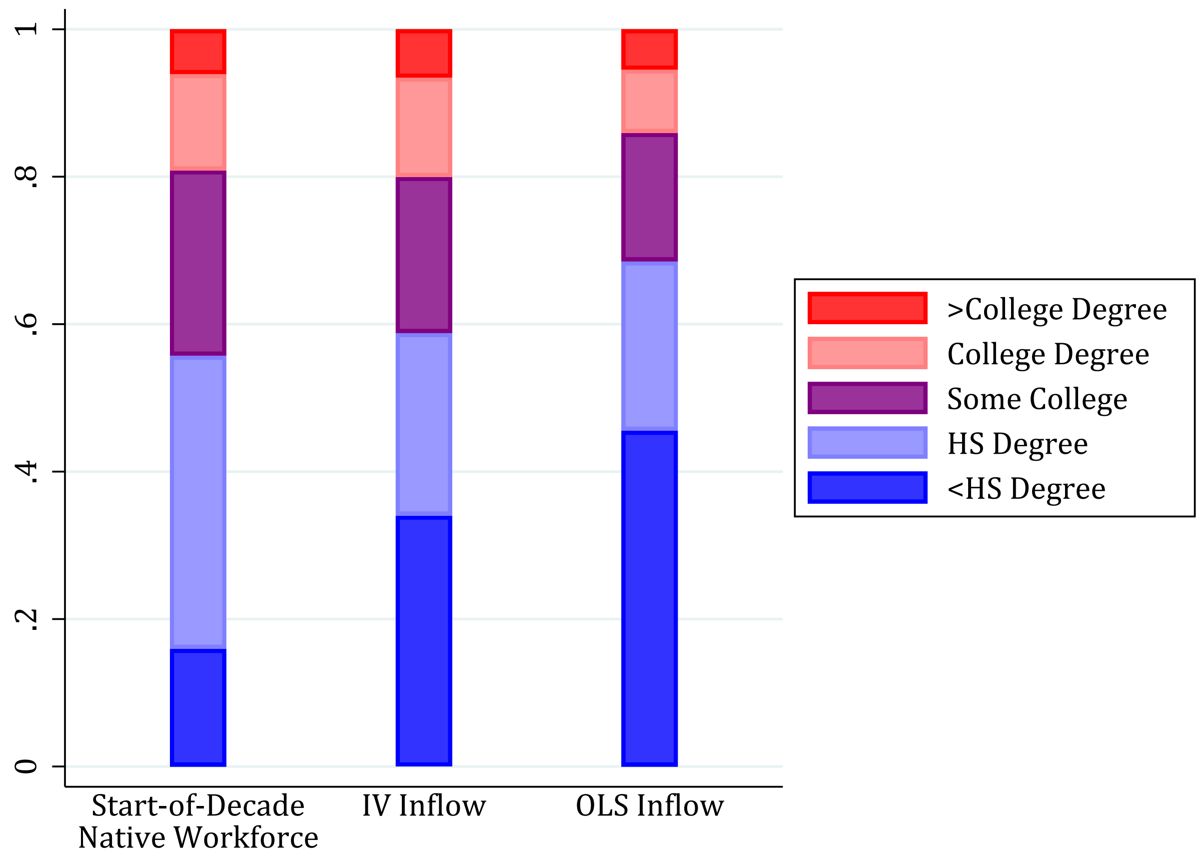
Figures and Tables

Figure 1: Instrument Balance Tests



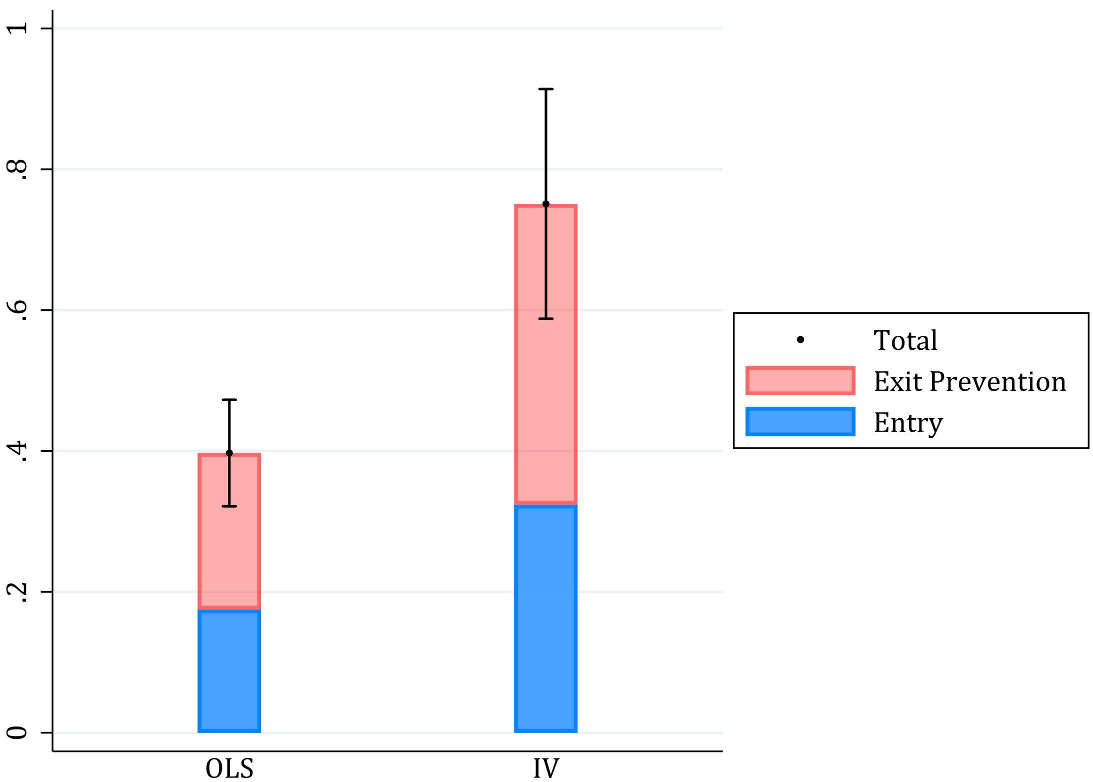
Notes: See Equation (2.4) for specification. Each specification is estimated using 88,806 observations that represent 722 CZ × 41 industries = 29,602 local industries observed for three decades, weighted by 1980 workforce size in the local industry. Each outcome is a standardized version of the indicated variable, regressed on the instrumental variable, $\Delta z_{gkt}^{\text{Emigrants}}$, along with commuting-zone-by-year and region-by-year-by-industry fixed effects. No additional controls are included in these specifications. "Balance outcomes" test whether the instrument is related to third factors that relate to or pre-study-period trends in our outcomes of interest. "True outcomes"—standardized versions of outcomes of interest in this paper—are presented for comparison. The group of control variables presented at the bottom of the figure are those that are included in most of the models estimated using Equation (2.1).

Figure 2: Educational Attainment of Immigrant Inflows and Receiving, Native Population (1980–2010)



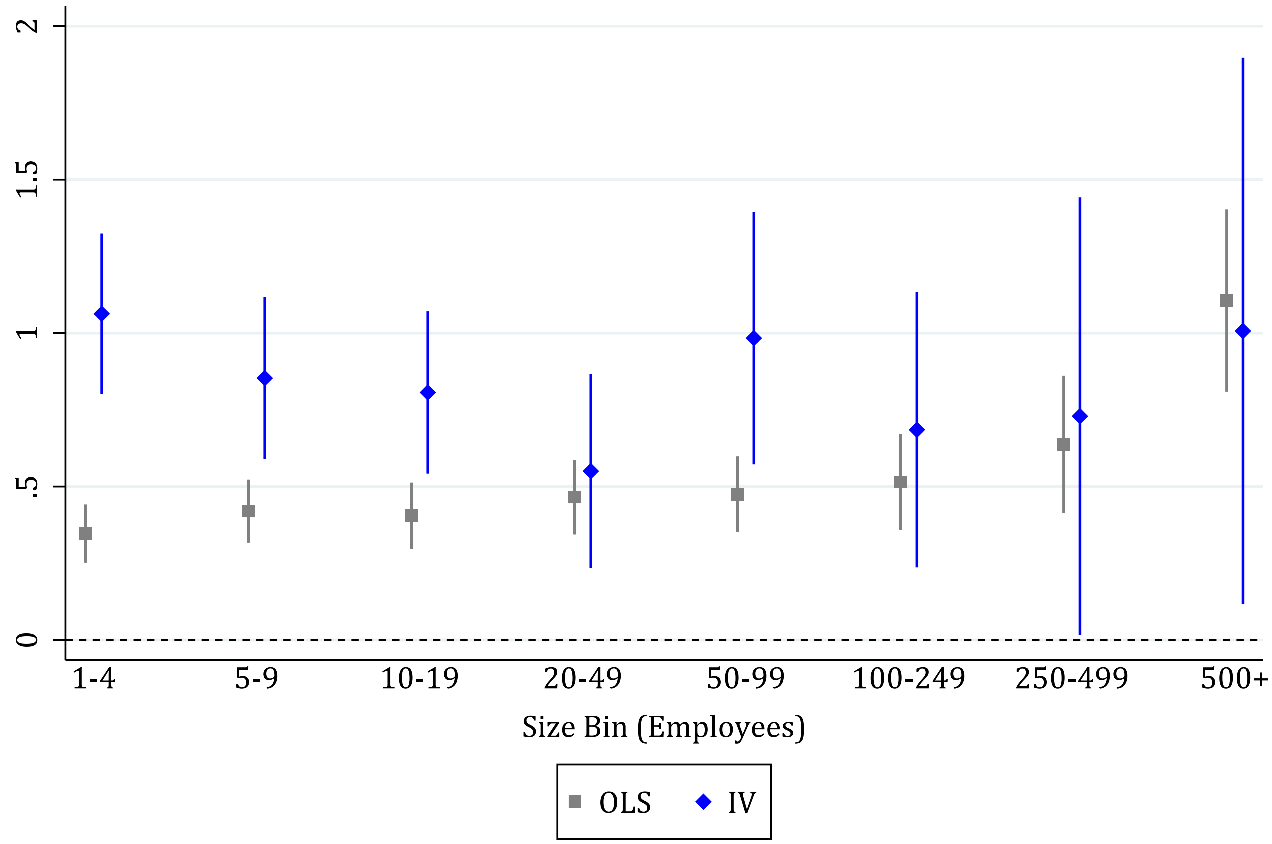
Notes: "Start-of-Decade Native Workforce" (left bar) constructed from publicly-available, U.S. Census Bureau demographic data (accessed through IPUMS-USA, Ruggles et al. (2019)). Each component of the "Start-of-Period Native Workforce" bar is generated by averaging the proportion of native workers in a that are from a given education grouping across local industries for 1980, 1990, and 2000. Each component of the "Inflow" bars (middle and right) obtained from estimating Equation (2.1) with $\Delta I_{gkt}^{\text{Educ. Level}}$ as the outcome, for mutually exclusive and exhaustive Educ. Levels, over the three decades covered by the time period 1980–2010. Each specification is estimated using 88,806 observations that represent 722 CZ × 41 industries = 29,602 local industries observed for three decades, weighted by 1980 workforce size in the local industry. These specifications are conducted using restricted-access demographic data from the U.S. Census Bureau and include control variables for start-of-period college share, start-of-period self-employment share, and start-of-period under-40 share in the local industry, along with a Bartik labor demand control and shift-share controls for exposure to international trade. Each specification also includes commuting-zone-by-year and region-by-industry-by-year fixed effects. IV specifications also include a control variable that interacts the sum of instrument shares with year fixed effects.

Figure 3: Effect of Immigration on Establishment Entry and Exit in Local Industries (1980–2010)



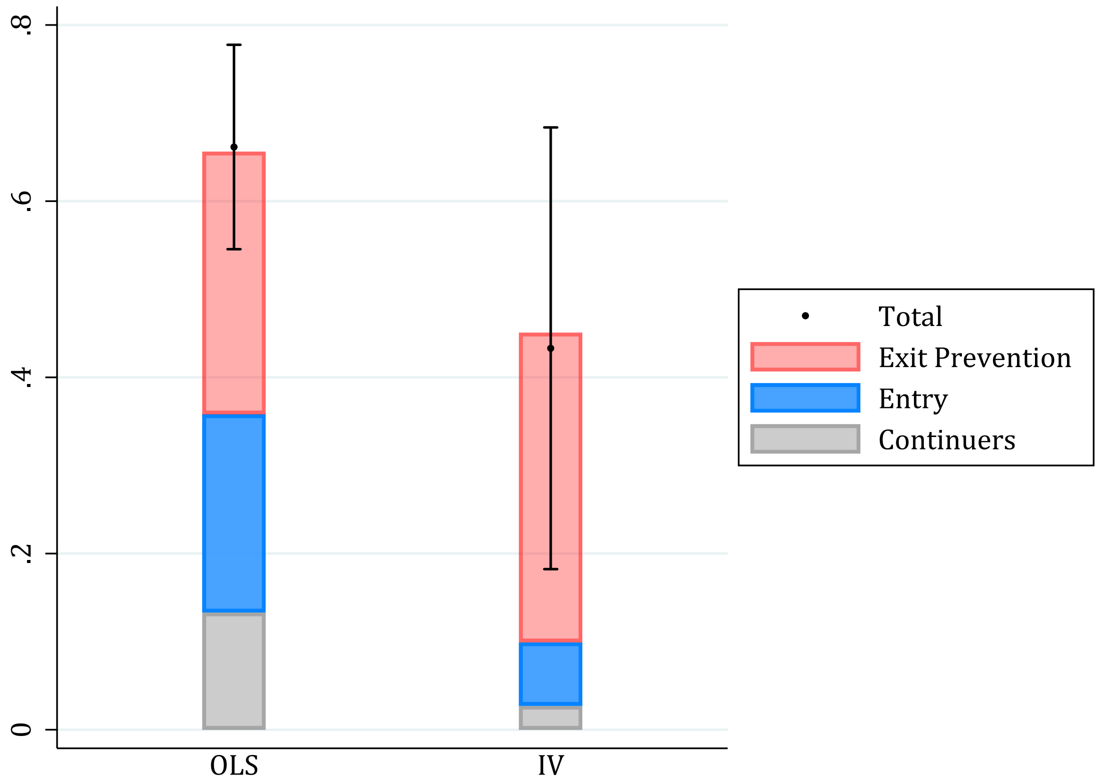
Notes: See Equation (2.1) for specification. Each bar adds up to the effect of a 1% immigration shock to a local industry's workforce on the DHS growth rate in the number of operating establishments in that local industry. This effect is also plotted using a black point, with capped spikes representing the 95% confidence interval around it. The contribution of establishment entry to this overall effect is in blue (bottom component) and the contribution of establishment exit prevention (negative establishment exit) is in red (top component). Each specification is estimated using 88,806 observations that represent 722 CZ × 41 industries = 29,602 local industries observed for three decades, weighted by 1980 workforce size in the local industry. Specifications include control variables for start-of-period college share, start-of-period self-employment share, and start-of-period under-40 share in the local industry, along with a Bartik labor demand control and shift-share controls for exposure to international trade. IV specification also includes a control variable that interacts the sum of instrument shares with year fixed effects. Each specification also includes commuting-zone-by-year and region-by-industry-by-year fixed effects. Appendix Table A3 contains the coefficients that underlie this figure and their corresponding standard errors.

Figure 4: Effect of Immigration on Establishment Presence in Local Industries, Within Size Bin (1980–2010)



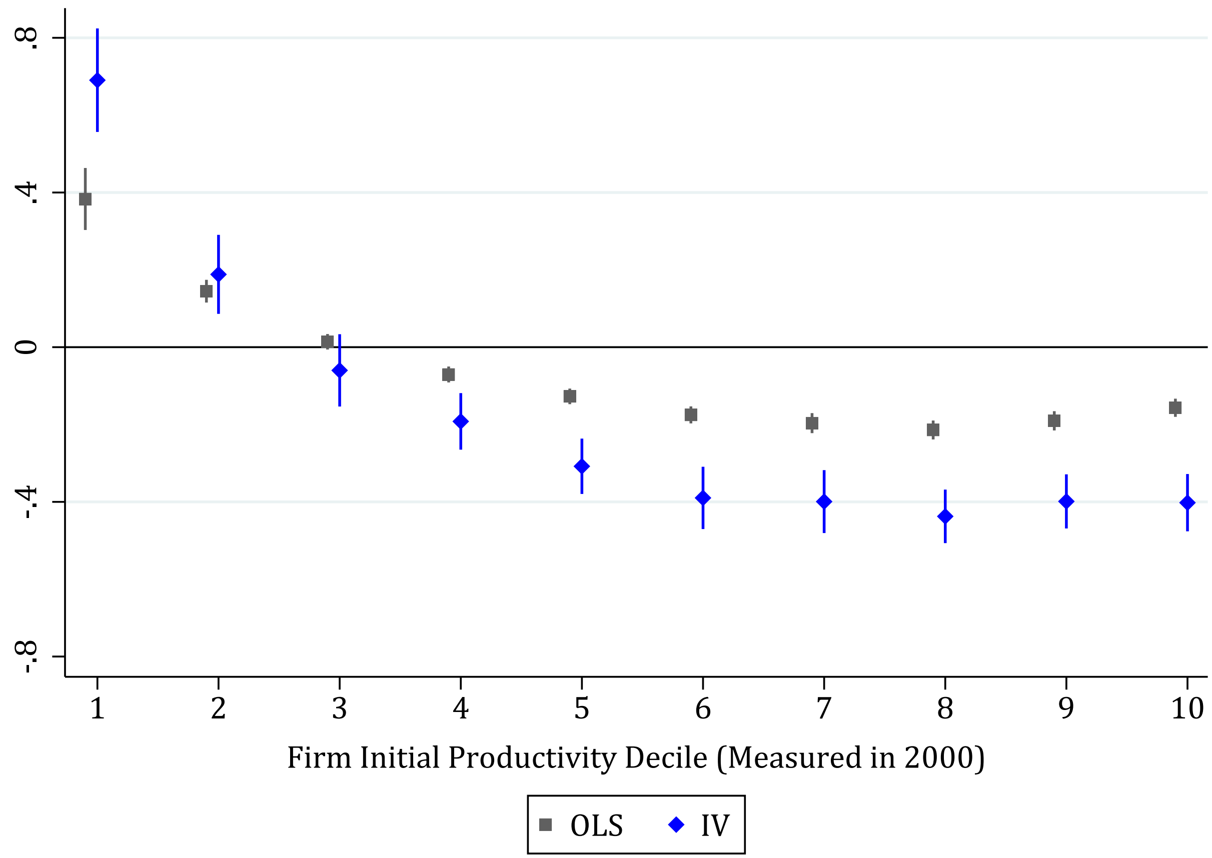
Notes: See Equation (2.1) for specification. Each coefficient represents the effect of a 1% immigration shock to a local industry's workforce on the within-size-bin DHS growth rate in the number of operating establishments in that local industry. Spikes around plotted coefficients represent 95% confidence intervals (standard errors clustered at the local industry level). Each specification is estimated using 88,806 observations that represent 722 CZ × 41 industries = 29,602 local industries observed for three decades, weighted by 1980 workforce size in the local industry. Specifications that generate coefficients include control variables for start-of-period college share, start-of-period self-employment share, and start-of-period under-40 share in the local industry, along with a Bartik labor demand control and shift-share controls for exposure to international trade. Each specification also includes commuting-zone-by-year and region-by-industry-by-year fixed effects. Finally, each specification also includes a control variable that interacts the sum of instrument shares with year fixed effects.

Figure 5: Decomposition of Immigrant-Induced Net Job Creation in Local Industries (1980–2010)



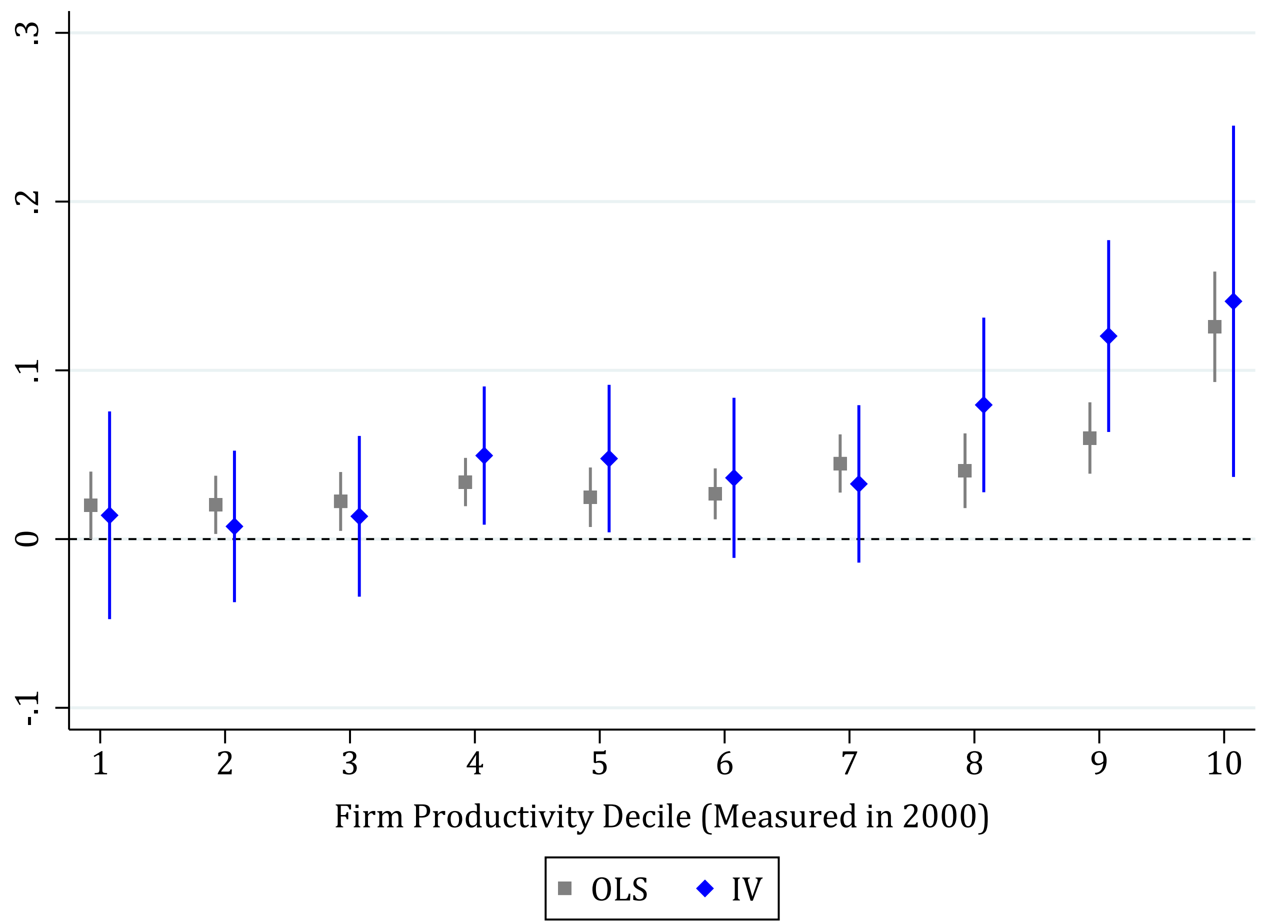
Notes: See Equation (2.1) for specification. Each bar adds up to the effect of a 1% immigration shock to a local industry's workforce on the DHS growth rate in employment minus a residual. The overall effect (including the residual) is also plotted using a black dot, with capped spikes representing the 95% confidence interval around it. The contribution of continuing establishments to this overall effect is in gray (bottom component). The contribution of establishment entry to this overall effect is in blue (middle component) and the contribution of establishment exit prevention (negative establishment exit) is in red (top component). Each specification is estimated using 88,806 observations that represent 722 CZ $\times$ 41 industries = 29,602 local industries observed for three decades, weighted by 1980 workforce size in the local industry. Specifications include control variables for start-of-period college share, start-of-period self-employment share, and start-of-period under-40 share in the local industry, along with a Bartik labor demand control and shift-share controls for exposure to international trade. Each specification alson includes commuting-zone-by-year and region-by-industry-by-year fixed effects. IV specification also includes a control variable that interacts the sum of instrument shares with year fixed effects. Appendix Table A4 contains the coefficients that underlie this figure and their corresponding standard errors.

Figure 6: Effect of Immigration on Establishment Exit, Stratifed by Initial Firm Productivity (2000–2015)



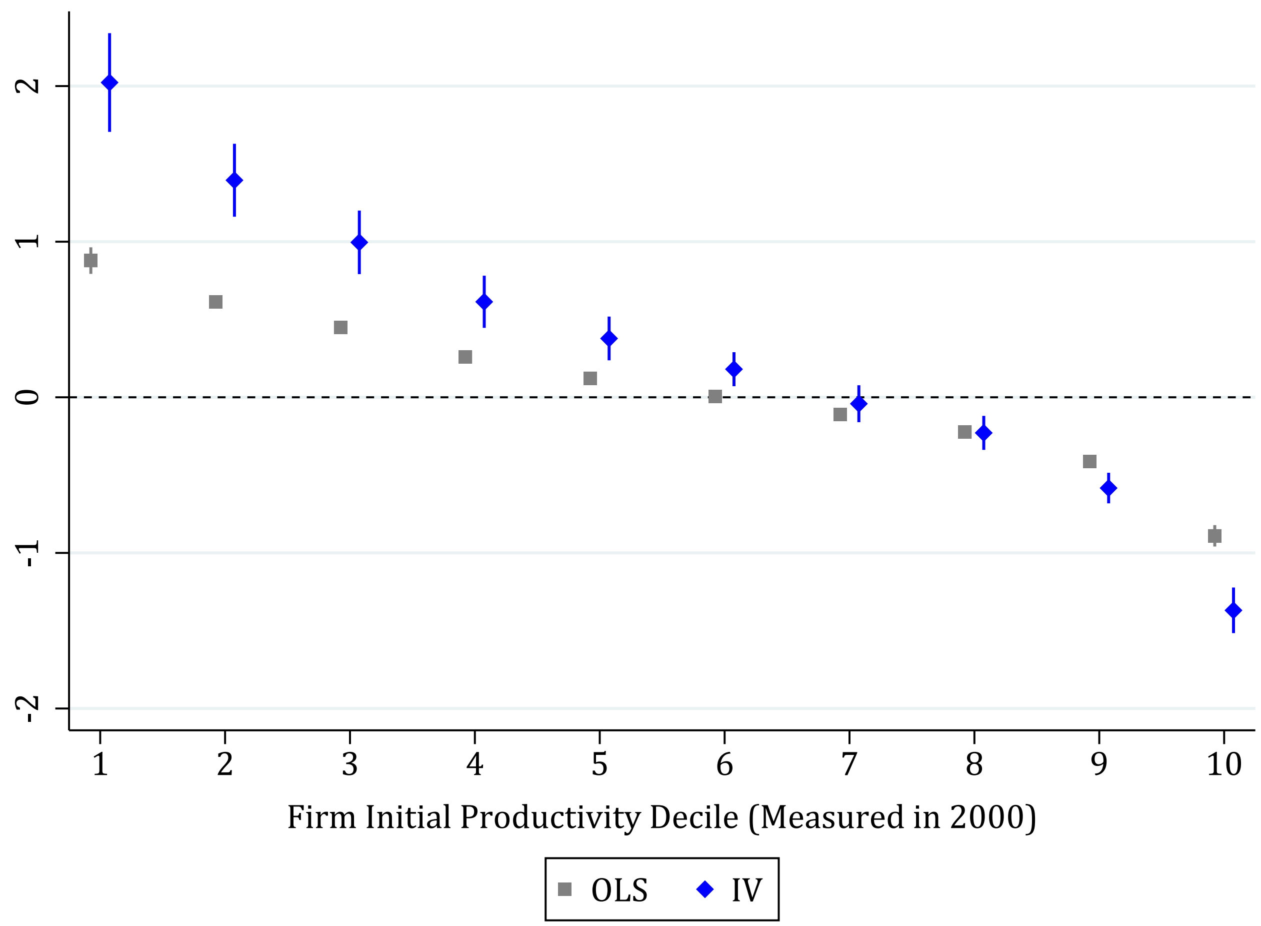
Notes: See Equation (4.2) for specification. Each coefficient represents the effect (in percent) of a 1% immigration shock to a local industry's workforce on the probability that a given establishment is no longer operating. An establishment is classified as "operating" if it has positive payroll and employment. Establishments are split into deciles based on their parent firm's national rank in revenues per worker within 5-digit NAICS code and age bin in 2000. Spikes indicate 95% confidence intervals (standard errors clustered at the local industry level). Each specification covers 4,739,000 establishments (rounded to avoid disclosure concerns), followed every five years until 2015. The 1st Stage F Statistic for the IV specification is 8.91. Each specification is estimated using 88,806 observations that represent 722 CZ × 41 industries = 29,602 local industries observed for three decades, weighted by 1980 workforce size in the local industry. Each specification includes control variables for 2000 college share, 2000 self-employment share, and 2000 under-40 share in the local industry, along with a Bartik labor demand control and shift-share controls for exposure to international trade. Each specification also includes establishment fixed effects, commuting-zone-by-year fixed effects, and region-by-industry-by-year fixed effects. IV specification also includes a control variable that interacts the sum of instrument shares with year fixed effects.

Figure 7: Effect of Immigration on Establishment Presence in Local Industries—Heterogeneity by Productivity (2000–2015)



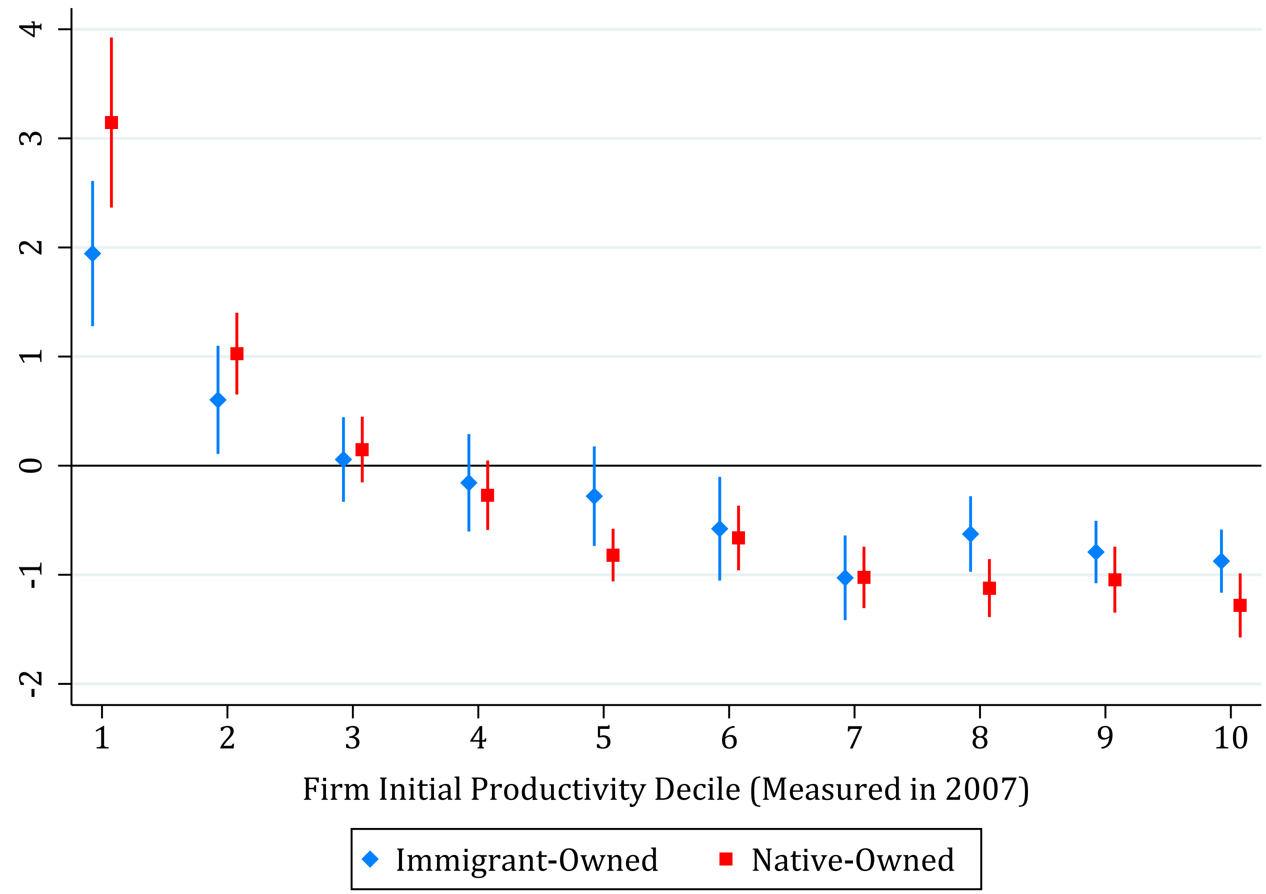
Notes: See Equation (4.4) for specification. Each coefficient represents the effect (in percent) of a 1% immigration shock to a local industry's workforce on the DHS growth rate in the number of operating establishments in that local industry whose parent firms have revenues per worker in the given productivity decile. Productivity is defined as revenues per worker and deciles are defined within 5-digit NAICS code by age bin. Spikes indicate 95% confidence intervals (standard errors clustered at the local industry level). There are 722 CZ $\times$ 41 industries = 29,602 local indsutry observations in each specification, weighted by 2000 workforce size. Each specification includes control variables for 2000 college share, 2000 self-employment share, and 2000 under-40 share in the local industry, along with a Bartik labor demand control and shift-share controls for exposure to international trade. Finally, each specification also includes commuting-zone fixed effects and region-by-industry fixed effects. IV specifications also include a control variable that interacts the sum of instrument shares with year fixed effects.

Figure 8: Effect of Immigration on Within-Establishment Log Average Earnings, Stratifed by Initial Firm Productivity (2000–2015)



Notes: See Equation (4.2) for specification. Each coefficient represents the effect (in percent) of a 1% immigration shock to a local industry's workforce on log payroll per worker at an establishment (conditional on survival). Establishments are split into deciles based on their parent firm's national rank in revenues per worker within 5-digit NAICS code and age group bin in 2000. Spikes indicate 95% confidence intervals (standard errors clustered at the local industry level). Each specification covers 4,739,000 establishments (rounded to avoid disclosure concerns), followed every five years until 2015. The 1st Stage F Statistic for the IV specification is 8.01. Each specification includes control variables for 2000 college share, 2000 self-employment share, and 2000 under-40 share in the local industry, along with a Bartik labor demand control and shift-share controls for exposure to international trade. Finally, each specification also includes establishment fixed effects, commuting-zone-by-year fixed effects, and region-by-industry-by-year fixed effects. IV specification also includes a control variable that interacts the sum of instrument shares with year fixed effects.

Figure 9: Effect of Immigration on Establishment Exit, Stratifed by Initial Firm Productivity and Ownership Nativity (IV Results, 2007–2017)



Notes: See Equation (5.11) for specification. Each coefficient represents the effect (in percent) of a 1% immigration shock to a local industry's workforce on the probability that a given establishment is no longer operating in a given year. An establishment is classified as "operating" if it has positive payroll and employment. Establishments are split into deciles based on their parent firm's national rank in sales per worker within 5-digit NAICS code and age bin in 2007. Ownership nativity also defined at the firm level in 2007. Spikes indicate 95% confidence intervals (standard errors clustered at the local industry level). Specification covers 773,000 establishments (rounded to avoid disclosure concerns), followed every five years until 2017, weighted by 2007 SBO survey weight. 1st Stage F Statistic is 10.19. Each specification includes control variables for 2007 college share, 2007 self-employment share, and 2007 under-40 share in the local industry, along with a Bartik labor demand control and shift-share controls for exposure to international trade. Finally, each specification also includes establishment fixed effects, commuting-zone-by-year fixed effects, and region-by-industry-by-year fixed effects. IV specification also includes a control variable that interacts the sum of instrument shares with year fixed effects. See Appendix Figure A1 for corresponding results estimated using OLS.

Table 1: Effect of Immigration on Establishment Presence (1980–2010)

Outcome:	ΔI_{gkt}	DHS Growth Rate in Establishment Count	
	(1)	(2)	(3)
$\Delta z_{gkt}^{\text{Emigrants}}$: Emigrants Instrument	0.2301***		
	(0.0218)		
ΔI_{gkt}: Immigrant Inflows per Initial Worker		0.3972***	0.7509***
		(0.0386)	(0.0832)
Instrument	N/A: 1st Stage	N/A: OLS	$\Delta z_{gkt}^{\text{Emigrants}}$
Observations	88,806	88,806	88,806

Notes: See Equation (2.1) for specification. Standard errors, clustered at local industry level, in parentheses. 88,806 observations represent 722 CZ × 41 industries = 29,602 local industries observed for three decades. In all specifications, observations are weighted by 1980 local industry workforce size. All specifications also include control variables for start-of-period college share, start-of-period self-employment share, and start-of-period under-40 share in the local industry, along with a Bartik labor demand control and shift-share controls for exposure to international trade. Each specification includes commuting-zone-by-year and region-by-industry-by-year fixed effects. IV specification also includes a control variable that interacts the sum of instrument shares with year fixed effects. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table 2: The Effect of Immigration on Establishment Presence—Stability and Robustness of IV Estimates (1980–2010)

Outcome:	DHS Growth Rate in Establishment Count				$\Delta\log(\text{Estabs})$
	(1)	(2)	(3)	(4)	(5)
ΔI_{gkt}: Immigrant Inflows per Initial Worker	1.284***	0.7675***	0.7509***	0.7852***	0.7741***
	(0.1264)	(0.0768)	(0.0832)	(0.0797)	(0.0835)
Native Inflows per Initial Worker				0.1188***	
				(0.0068)	
Instrument	$\Delta z_{gkt}^{\text{Emigrants}}$	$\Delta z_{gkt}^{\text{Emigrants}}$	$\Delta z_{gkt}^{\text{Emigrants}}$	$\Delta z_{gkt}^{\text{Emigrants}}$	$\Delta z_{gkt}^{\text{Emigrants}}$
Time Period	1980–2010	1980–2010	1980–2010	1980–2010	1980–2010
1st Stage F Statistic	216.7	125.6	111.4	111.1	111.4
Within R^2	0.0557	0.0051	0.0033	0.0222	0.0043
Year FE	✓	(r)	(r)	(r)	(r)
Commuting Zone × Year FE		✓	✓	✓	✓
Region × Industry Group × Year FE		✓	✓	✓	✓
Controls			✓	✓	✓
Observations	88,806	88,806	88,806	88,806	88,806

Notes: See Equation (2.1) for specification. Standard errors, clustered at local industry level, in parentheses. Each decade studied contains 722 CZ × 41 industries = 29,602 local industries. In all specifications, observations are weighted by 1980 local industry workforce size. Where indicated, specifications also include control variables for start-of-period college share, start-of-period self-employment share, and start-of-period under-40 share in the local industry, along with a Bartik labor demand control and shift-share controls for exposure to international trade. Each specification also include a control variable that interacts the sum of instrument shares with year fixed effects. (r) indicates that a fixed effect is redundant based on other fixed effect sets contained in the given model. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table 3: Effect of Immigration on Establishment-Level Outcomes, 2000–2015

Outcome:	$\mathbb{1}\{\text{Not Operating}\}_{et}$ (1)	$\log(\text{Employment})_{et}$ (2)	$\log(\text{Payroll p.w.})_{et}$ (3)	$\log(\text{Revenues})_{f(e)t}$ (4)	$\log(\text{Revenes p.w.})_{f(e)t}$ (5)
Panel A: OLS					
I_{gkt}: Immigrant Workers per 2000 Workforce	-0.0369***	0.0658***	0.0200***	0.1162***	0.0436***
	(0.0051)	(0.0090)	(0.0076)	(0.0196)	(0.0160)
Panel B: Emigrants IV ($z_{gkt}^{\textbf{Emigrants}}$)					
I_{gkt}: Immigrant Workers per 2000 Workforce	-0.2139***	0.0749	0.0309	0.1260	0.0034
	(0.0344)	(0.0575)	(0.0609)	(0.1506)	(0.1110)
Outcome Measurement Level	Establishment	Establishment	Establishment	Firm	Firm
Establishments	6,180,000	6,180,000	6,180,000	4,739,000	4,739,000

Notes: See Equation (4.1) for specification. Standard errors, clustered at local industry level, in parentheses. Each establishment followed for four observations ($t \in \{2000, 2005, 2010, 2015\}$). Establishment and observation counts rounded to avoid disclosure concerns. With the exception of $\mathbb{1}\{\text{Not Operating}\}_{et}$, outcomes are missing when establishment is not in operation. 1st Stage F statistic for Panel B is 66.1. All specifications include control variables for 2000 college share, 2000 self-employment share, and 2000 under-40 share in the local industry, along with a Bartik labor demand control and shift-share controls for exposure to international trade. Each specification also includes establishment fixed effects, commuting-zone-by-year fixed effects, and region-by-industry-by-year fixed effects. IV specifications also include a control variable that interacts the sum of instrument shares with year fixed effects. "p.w." stands for per worker. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table 4: The Effect of Immigration on Stratified Establishment Exit—Robustness (IV Results, 2000–2015)

	Outcome: $\mathbb{1}\{\text{Not Operating}\}_{et}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
I_{gkt}: Immigrant Workers per 2000 Workforce ($\hat{\eta}_0$)	0.7397***	0.7790***	1.135***	0.6191***	0.9457***	0.9268***	1.195***
	(0.0729)	(0.0808)	(0.0896)	(0.0603)	(0.1024)	(0.071)	(0.0966)
$I_{gkt} \times \left[\text{Prod. Pctl./100}\right]_e$ ($\hat{\eta}_1$)	-3.364***	-3.454***	-4.690***	-2.891***	-3.593***	-2.733***	-5.118***
	(0.2170)	(0.2496)	(0.2844)	(0.1739)	(0.2837)	(0.1608)	(0.3516)
$I_{gkt} \times \left[\text{Prod. Pctl./100}\right]_e^2$ ($\hat{\eta}_2$)	2.330***	2.341***	2.923***	1.673***	2.035***	0.707***	3.473***
	(0.1588)	(0.1799)	(0.1940)	(0.118)	(0.1887)	(0.1059)	(0.2582)
N_{gkt}: Native Workers per 2000 Workforce		-0.0306***					
		(0.0076)					
$N_{gkt} \times \left[\text{Prod. Pctl./100}\right]_e$		0.0533*					
		(0.0292)					
$N_{gkt} \times \left[\text{Prod. Pctl./100}\right]_e^2$		0.0080					
		(0.0225)					
Implied Crossing Point Pctl.	27.1	27.8	29.7	25.0	32.2	37.6	29.1
Productivity Measure	Revenues p.w.	Revenues p.w.	Revenues	Employment	Payroll p.w.	Employment	Payroll p.w.
Productivity Measurement Level	Firm	Firm	Firm	Firm	Firm	Estab.	Estab.
Productivity Rank Bin	Age Group × NAICS-5D	Age Group × NAICS-5D	Age Group × NAICS-5D	Age Group × NAICS-5D	Age Group × NAICS-5D	Local Industry	Local Industry
1st Stage F Statistic	29.72	29.66	30.42	21.90	21.36	22.04	22.04
Establishments	4,739,000	4,739,000	4,739,000	6,180,000	6,180,000	6,180,000	6,180,000

Notes: See Equation (4.3) for specification. Standard errors, clustered at local industry level, in parentheses. Each establishment followed for four observations ($t \in \{2000, 2005, 2010, 2015\}$). Establishment counts rounded to avoid disclosure concerns. "Implied Crossing Point Pctl." refers to the rank percentile of productivity that the effect of I_{gkt} on $\mathbb{1}\{\text{Not Operating}\}_{et}$ turns from positive to negative. Percentiles (Pctl.) are determined based on a unit's rank of "Productivity Measure" within "Productivity Rank Bin," where a unit corresponds to the "Productivity Measurement Level" row. "p.w." stands for per worker. All specifications include control variables for 2000 college share, 2000 self-employment share, and 2000 under-40 share in the local industry, along with a Bartik labor demand control and shift-share controls for exposure to international trade. Each specification also includes establishment fixed effects, commuting-zone-by-year fixed effects, and region-by-industry-by-year fixed effects. Each specification also includes a control variable that interacts the sum of instrument shares with year fixed effects. Vector of instrumental variables is $\left(z_{gkt}^{\text{Emigrants}}, z_{gkt}^{\text{Emigrants}} \times \left[\text{Prod. Pctl./100}\right]_e, z_{gkt}^{\text{Emigrants}} \times \left[\text{Prod. Pctl./100}\right]_e^2\right)$. See Table A5 for corresponding OLS-estimated results. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

A Data

A.1 Business Data

The U.S. Census Bureau's Longitudinal Business Database (LBD) is constructed from administrative tax records for each U.S. non-farm, employee-hiring, private-sector establishment. The establishments considered in this paper are those that can be assigned a firm identifier, those that can be assigned a Fort-Klimek consistent NAICS code (Fort and Klimek, 2018), those whose consistent NAICS code is then among those covered in Table A1, and those in the contiguous United States (including Washington D.C., but excluding Alaska and Hawaii). This comprises the vast majority of the private U.S. economy. Most of the analyses in this paper use the original LBD, which covers the years 1976 through 2016. However, the analysis in Section 5.5.2 uses the new, Revised LBD, which contains 2017 and 2018 as well. In any given analysis, establishments are assigned their earliest possible industry group in order to avoid misclassifying an industry switch as an establishment exit.

A.2 Demographic Data

Immigrant exposure variables and several control variables are measured using restricted-access U.S. Census Bureau demographic data from the 1980, 1990, and 2000 Long-Form Decennial Censuses and the 2005 through 2019 American Community Surveys (ACS). The underlying sample of respondents that I use to construct these measures consists of employed workers (self-employed or employees) who can be assigned a country of origin, who reside in the contiguous United States (including Washington, D.C., but excluding Alaska and Hawaii), and who work in an industry group from Table A1. Immigrant workers are defined as those who indicated that they are either a naturalized citizen or a non-citizen. All other workers are defined as native. Population estimates are generated by summing over survey weights from the underlying sample.

The ACS is a yearly survey that contains smaller underlying sample sizes than the Decennial Census Long Form. Given the level of detail in the unit of analysis—commuting zone by industry group—I average across ACS years to increase the underlying sample size. Specifically, for all years other than 2005, measures from the ACS are averaged over five years to eliminate noise and keep underlying sample sizes similar to measures from the Decennial Censuses. For example, estimates of immigrant presence in a local industry in 2010 is the average of immigrant presence in that local industry over the period 2008-2012. This cannot be done for 2005 because the 2001-2004 ACS were experimental products that were not representative at sub-state levels.

A.3 Emigration Data

I use the Brain-Drain Data (Brücker et al., 2013)—specifically the "Migration by Gender" dataset—to construct $\Delta z_{gkt}^{\text{Emigrants}}$, the instrumental variable for analyses in Section 3. These data contain total emigration stocks in origin-destination pairs for 20 OECD destinations, including the U.S. Stocks are mostly obtained from destination country microdata, and are sometimes imputed, as described in Brücker et al. (2013).

The IAB data covers 1980 but does not cover 2015. Thus, I use the United Nations Population Division's (UNPD) International Migration Stock 2019—which does not cover 1980—to construct $z_{gkt}^{\text{Emigrants}}$, the instrumental variable for analyses in Sections 4 and 5.5.1 and $z_{gk,t-2}^{\text{Emigrants}}$, the instrumental variable for analyses in Section 5.5.2. These data also include origin-destination pair emigration stocks constructed mostly from destination country microdata, but for a far larger set of destination countries.

In order to keep the instrumental variable as consistent as possible, I always sum emigrants in non-U.S. destinations for 18 OECD member nations: Australia, Austria, Canada, Chile, Denmark, Finland, France, Germany, Greece, Ireland, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom. Other than the U.S., Luxembourg is dropped from the destination list because it is combined with Belgium as an origin country in the data used to construct the control variables for trade disclosure (described in Appendix Section A.5). This takes us from the original 20 destinations in the IAB data to 18. All destination countries are dropped as potential origins. In particular, this removes many Europe-to-Europe moves that started to occur after the introduction of the Eurozone in 1999. Results are robust to dropping Canada as a destination country, prompted by concerns of destination substitution between the U.S. and Canada (available upon request).

121 origin countries are covered. These generally represent the largest 121 migration origin countries, but some aggregations are made in order to account for changing boundaries over time during the study period and aggregations that come with the IAB and UNPD data. These include combining countries from the former Soviet Union, countries from the former Czechoslovakia, countries from the former Yugoslavia, South Sudan and Sudan, disputed and undisputed territories associated with China (including Taiwan), Eritrea and Ethiopia, and Israel and Palestinian/disputed territories.

A.4 Industry Classifications and Summary Statistics

Because industry classifications differ both across Census years and between the Census and NAICS, constructing the industry groups involved multiple steps. I first use the 1990 Decennial Census industry codes as a bridge between different Census industry classification systems, as is done in IPUMS-USA (Ruggles et al., 2019).[56] I then construct a crosswalk between the 1990 Decennial Census industry codes and 3-digit 2012 NAICS codes, which available for all years in the LBD from Fort and Klimek (2018). In some cases, the 1990 industry classification corresponds to more than one 3-digit NAICS code, and in some cases a 3-digit NAICS code corresponds to more than one 1990 industry classification. The industry groups I use therefore generally represent the smallest possible mutually-exclusive sets of industry classifications.

For example, 1990 Census Industry classification code 132 is "Knitting mills" and corresponds to NAICS Codes 315 and 313. However, NAICS code 313 also covers "Yarn, thread, and fabric mills," which is 1990 Census Industry classification code 142. Additionally, NAICS code 315 also includes manufacturing of "Apparel and accessories other than knitting." Manufacturing of apparel and accessories, knitting mills, and yarn, thread, and fabric mills are therefore all covered in the same industry grouping in my analysis.

Some additional aggregations are made to ensure that industry groups do not vary excessively in size. The Agriculture (NAICS 11), Mining (NAICS 21), and Public (NAICS 92) sectors along with the Postal Service (NAICS 491), Fund, Trusts, and Other Financial Vehicles (NAICS 525), and Private Households (NAICS 814) are dropped from the analysis due to relatively less reliable coverage in the LBD. The final set of industry groups can be seen in Table A1.

[56] Crosswalks provided by IPUMS-USA between the 1990 and other Census year classifications, as well as between the 1990 Census industry classifications and NAICS codes, were crucial to this process.

Table A1: Industry Groups

Industry Group	1990 Census Codes	2007 NAICS Codes	Worker Educ. Designation	Tradability Designation
Construction	60	23	High-School Equivalent	Non-Tradable
Management of companies	710	55	College Equivalent	Tradable
Utilities	422, 450, 451, 470–472	22, 486, 562	High-School Equivalent	Non-Tradable
Manufacturing – Food	100, 101, 102, 110 , 111, 112, 120, 121, 122, 130, 610	311–312	High-School Equivalent	Tradable
Manufacturing – Clothing	132, 140, 142, 150, 151, 152, 220, 221, 222	313–316	High-School Equivalent	Tradable
Manufacturing – Wood & Furniture +	160–162, 231, 232, 241, 242, 250–252, 261, 262	321, 322, 327, 337	High-School Equivalent	Tradable
Manufacturing – Plastics +	180–182, 190–192, 200, 201, 210–212	324–326	College Equivalent	Tradable
Manufacturing – Metals & Machinery	270–272, 280–282, 290–292, 300, 301, 310–312, 320, 321, 331, 332, 380	331–333	High-School Equivalent	Tradable
Manufacturing – Electrical & Household	322, 340–342, 350, 371, 372, 381, 390, 391	334, 335, 339	College Equivalent	Tradable
Manufacturing – Transportation	351, 352, 360–362, 370	336	High-School Equivalent	Tradable
Printing & Publishing	171, 172	323, 511	College-Equivalent	Tradable
Wholesale Trade – Durable	500, 501, 502, 510–512, 521, 530–532	423	College Equivalent	Tradable
Wholesale Trade – Nondurable	540–542, 550–552, 560–562	424	High-School Equivalent	Tradable
Retail Trade – Vehicles	612, 620, 622	441	High-School Equivalent	Non-Tradable
Retail Trade – Household Durables	580–582, 631–633	442–444	College Equivalent	Non-Tradable
Retail Trade – Food & Gas	601, 602, 611, 621, 650	445, 447	High-School Equivalent	Non-Tradable
Retail Trade – Misc.	590, 640, 642, 651, 652, 661, 662, 681, 682	446, 451, 453	College Equivalent	Non-Tradable
Retail Trade – Apparel	623, 630, 660	448	High-School Equivalent	Non-Tradable
Retail Trade – Dept. & Variety Stores	591, 592, 600	452	High-School Equivalent	Non-Tradable
Retail Trade – Fuel, Catalog, Vending	663, 670–672	454	High-School Equivalent	Non-Tradable
Misc. Transportation	400, 401, 402, 420, 421	481–483, 485	High-School Equivalent	Tradable
Trucking	410	484, 492	High-School Equivalent	Non-Tradable
Warehousing & Storage	411	493	High-School Equivalent	Tradable
Non-Telephone Communication	440, 852	515, 519	College Equivalent	Non-Tradable
Telecomm & Data Processing	441, 442, 732	517, 518	College Equivalent	Non-Tradable
Savings Institutions	700–702	521, 522	College Equivalent	Non-Tradable
Insurance	711	524	College Equivalent	Tradable
Real Estate	712	531	College Equivalent	Non-Tradable
Professional Services	12, 721, 741, 841, 882, 890–893	541, 711	College Equivalent	Tradable
Admin. & Support Services	20, 432, 722, 731, 740	561	College Equivalent	Non-Tradable
Educational Services	842, 850, 851, 860	611	College Equivalent	Non-Tradable
Health Services excl. Hospitals	812, 820–822, 830, 840	621	College Equivalent	Non-Tradable
Hospitals	831	622	College Equivalent	Non-Tradable
Nursing & Residential Care Facilities	832, 870	623	High-School Equivalent	Non-Tradable
Social Services	861–863	624	College Equivalent	Non-Tradable
Entertainment Services	742, 800–802, 810, 872	512, 532, 712, 713	College Equivalent	Non-Tradable
Lodging	762, 770	721	High-School Equivalent	Tradable
Eating & Drinking Places	641	722	High-School Equivalent	Non-Tradable
Repair Services	750–752, 760, 782, 790	811	High-School Equivalent	Non-Tradable
Personal Services	771, 772, 780, 781, 791	812	High-School Equivalent	Non-Tradable
Unions & Religious Organizations	873, 880, 881	813	College Equivalent	Non-Tradable

Table A2: Summary Statistics (Publicly-Available Data)

	1980 Workforce	1980–2010 ΔI_{gkt}: Immigrant Inflows per Initial Worker		1980–2010 Δ Workforce per Initial Worker		2000–2010 DHS Growth Rate in Establishment Count	
Industry Group	Mean	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Admin. & Support Services	28,412	0.143	0.162	0.612	0.604	0.100	0.108
Construction	51,145	0.070	0.109	0.232	0.347	-0.089	0.132
Eating & Drinking Places	48,400	0.075	0.094	0.394	0.333	0.174	0.113
Educational Services	28,537	0.027	0.062	0.254	0.344	0.254	0.165
Entertainment Services	24,966	0.044	0.072	0.412	0.605	0.025	0.082
Health Services excl. Hospitals	23,807	0.065	0.093	0.629	0.352	0.194	0.115
Hospitals	48,735	0.036	0.055	0.297	0.297	-0.069	0.191
Insurance	33,594	0.012	0.034	0.154	0.300	0.011	0.092
Lodging	11,515	0.077	0.114	0.230	0.435	0.055	0.149
Management of companies	25,535	0.055	0.089	0.489	0.934	0.004	0.169
Manufacturing, Clothing	40,229	-0.020	0.137	-0.268	0.457	-0.615	0.268
Manufacturing, Electrical & Household	78,275	0.025	0.075	0.001	0.404	-0.170	0.140
Manufacturing, Food	16,349	0.045	0.099	-0.015	0.314	-0.021	0.190
Manufacturing, Metals & Machinery	74,356	-0.002	0.046	-0.117	0.280	-0.147	0.128
Manufacturing, Plastics +	36,251	0.023	0.074	0.018	0.386	0.005	0.168
Manufacturing, Transportation	101,875	0.016	0.060	0.053	0.653	-0.430	0.232
Manufacturing, Wood & Furniture +	20,281	0.015	0.059	-0.052	0.353	-0.148	0.136
Misc. Transportation Transportation	29,749	0.024	0.058	-0.029	0.335	0.063	0.229
Non-Telephone Communication	4,352	0.038	0.087	0.396	0.682	-0.214	0.280
Nursing & Residential Care Facilities	9,714	0.056	0.097	0.460	0.431	0.223	0.187
Personal Services	15,327	0.057	0.104	0.114	0.330	0.015	0.146
Printing & Publishing	34,245	0.004	0.047	-0.013	0.302	-0.087	0.153
Professional Services	81,687	0.038	0.052	0.380	0.328	0.163	0.128
Real Estate	25,828	0.039	0.067	0.294	0.455	0.193	0.135
Repair Services	16,736	0.043	0.079	0.142	0.287	-0.102	0.089
Retail Trade, Household Durables	15,068	0.026	0.062	0.234	0.376	-0.105	0.099
Retail Trade, Apparel	15,837	0.018	0.080	0.095	0.369	-0.044	0.161
Retail Trade, Dept. & Variety Stores	27,785	0.020	0.056	0.080	0.263	0.175	0.135
Retail Trade, Food & Gas	32,599	0.031	0.059	0.102	0.214	-0.062	0.105
Retail Trade, Fuel, Catalog, Vending	6,847	0.025	0.071	0.095	0.576	0.202	0.207
Retail Trade, Misc.	19,697	0.032	0.059	0.284	0.274	-0.078	0.093
Retail Trade, Vehicles	11,246	0.021	0.052	0.124	0.223	-0.064	0.105
Savings Institutions	41,095	0.030	0.052	0.216	0.302	0.108	0.121
Social Services	3,743	0.106	0.173	0.699	0.673	0.207	0.133
Telecomm & Data Processing	28,661	0.059	0.085	0.267	0.411	-0.015	0.176
Trucking	13,974	0.040	0.075	0.236	0.331	-0.006	0.170
Unions & Religious Organizations	12,085	0.016	0.061	0.220	0.327	0.065	0.131
Utilities	7,219	0.025	0.051	0.172	0.348	0.312	0.238
Warehousing & Storage	2,840	0.086	0.203	0.444	3.252	0.644	0.302
Wholesale Trade, Durable	32,044	0.019	0.053	0.088	0.364	0.818	0.092
Wholesale Trade, Nondurable	25,090	0.029	0.076	-0.008	0.241	0.720	0.145
Total	43,955	0.035	0.084	0.162	0.454	0.017	0.300

Notes: Data obtained from IPUMS-USA (Ruggles et al., 2019) and County Business Patterns. All statistics weighted by 1980 workers in local industry.

A.5 Construction of Control Variables

A.5.1 Start-of-Period Shares

Specifications with control variables include controls for "start-of-period" college share, self-employment share, and under-40 share in the population. These are self-explanatory: the proportion of employed workers that have a college degree, are self-employed, and are under 40 years-old, respectively. Start-of-period is defined as the start of the decade for each of the decades studied in Section 3 using Equation (2.1) (in this case, these controls update over time). For analyses in Section 4 and 5.5.1, start-of-period is fixed in 2000 (in this case, these controls do not update over time). For analyses in Section 5.5.2, start-of-period is fixed in 2007.

A.5.2 Bartik Labor Demand Control

The structure of the control variable for labor demand mimics the instrumental variable for labor demand proposed by Bartik (1991). It is included because this paper seeks to isolate labor demand *responses* to labor supply shocks from immigration as opposed to labor demand shocks that are generated by the same factor that induces the immigration labor supply shocks. The control variable takes advantage of the fact that the LBD data contains consistent 5-digit NAICS codes over time, whereas my industry groupings are aggregations of 3-digit NAICS codes. Specifically, letting k' denote a 5-digit NAICS code and k denoting an industry group as usual,

$$\text{Bartik}_{gkt} = \sum_{k' \in k} \left[\left(\frac{\text{Employment}_{gk',1980}}{\sum_g \text{Employment}_{gk',1980}} \right) \times \sum_g \text{Employment}_{gk't} \right]$$

That is, the stock of national employment in 5-digit NAICS code at time t, measured using the LBD, is projected into local industries based on the proportion of national, 1980 5-digit NAICS code employment was located in commuting zone g to generate a predicted amount of employment in 5-digit NAICS code in commuting zone g at time t. These predictions are summed over k' to generate a predicted level of employment in local industry gk.

I use Bartik_{gkt} to control for labor demand shocks caused by the interaction between local industry specialization and secular trends in detailed industries. For example, an area specialized in a tradable manufacturing industry as of 1980 likely experienced large declines in labor demand due to import competition from China Autor et al. (2013).

Specifically, for analyses in Section 3 using Equation (2.1), I use the DHS growth rate in Bartik_{gkt}:

$$\frac{\text{Bartik}_{gkt} - \text{Bartik}_{gk,t-10}}{\left(\frac{\text{Bartik}_{gkt} + \text{Bartik}_{gk,t-10}}{2} \right)}$$

For analyses in Section 4.3, I use

$$\frac{\text{Bartik}_{gk,t=2015} - \text{Bartik}_{gk,t=2000}}{\left(\frac{\text{Bartik}_{gk,t=2015} + \text{Bartik}_{gk,t=2000}}{2} \right)}$$

For analyses in Section 4 and 5.5.1, I use a version in levels:

$$\frac{\text{Bartik}_{gkt}}{\left(\frac{\text{Bartik}_{gkt} + \text{Bartik}_{gk,t=2000}}{2} \right)}$$

Finally, for the analyses in Section 5.5.2, I use a similar version in levels:

$$\frac{\text{Bartik}_{gkt}}{\left(\frac{\text{Bartik}_{gkt}+\text{Bartik}_{gk,t=2007}}{2}\right)}$$

A.5.3 Controls for Trade Exposure

I also control for local industry trade exposure that may also arise from immigrant worker ties to their origin countries. These ties may confound results if origin-country shocks generate both pressure to emigrate and increased trade activity between these origin countries and the U.S. These controls variables use data from the World Bank's World Integrated Trade Solution, which contain data on real trade flows from the 121 aforementioned sending countries to and from the U.S. starting in 1980.

I use these data to construct shift-share control variables for trade exposure

$$[\text{Flow}]\ \text{Exposure}_{gkt} = \frac{1}{E_{gk,1980}} \sum_{o} \pi_{og,1980} \times \text{Prop. Traded}_{kt} \times [\text{Flow}]_{ot}$$

where $[\text{Flow}] \in \{\text{Imports}, \text{Exports}\}$ and Prop. Traded_{kt} is the proportion of industry k's workforce that is employed in a "Traded" 6-digit NAICS code industry according to the Porter classification system (Porter, 2003). $[\text{Flow}]\ \text{Exposure}_{gkt}$ is differenced over 10 and 15 years for the analyses in Section 3 and 4.3, respectively. It is included in levels in the analyses in Sections 4, 5.5.1, and 5.5.2.

B Supplemental Details and Results for Empirical Analyses

B.1 Additional Details for Main Tables and Figures

Table A3: The Effect of Immigration on Establishment Entry, Exit, and Overall Presence (Figure 3)

	DHS Growth Rate in Operating Establishments (1)	=	Contribution of: Entrants (2)	–	Exits (3)
Panel A: OLS					
ΔI_{gkt}: Immigrant Inflows per Initial Worker	0.3972***		0.1750***		-0.2209***
	(0.0386)		(0.0248)		(0.0224)
Percent of Total	100		44.06		55.61
Panel B: Emigrants IV ($z_{gkt}^{\textbf{Emigrants}}$)					
ΔI_{gkt}: Immigrant Inflows per Initial Worker	0.7509***		0.3239***		-0.4279***
	(0.0832)		(0.0710)		(0.0489)
Percent of Total	100		43.13		56.99

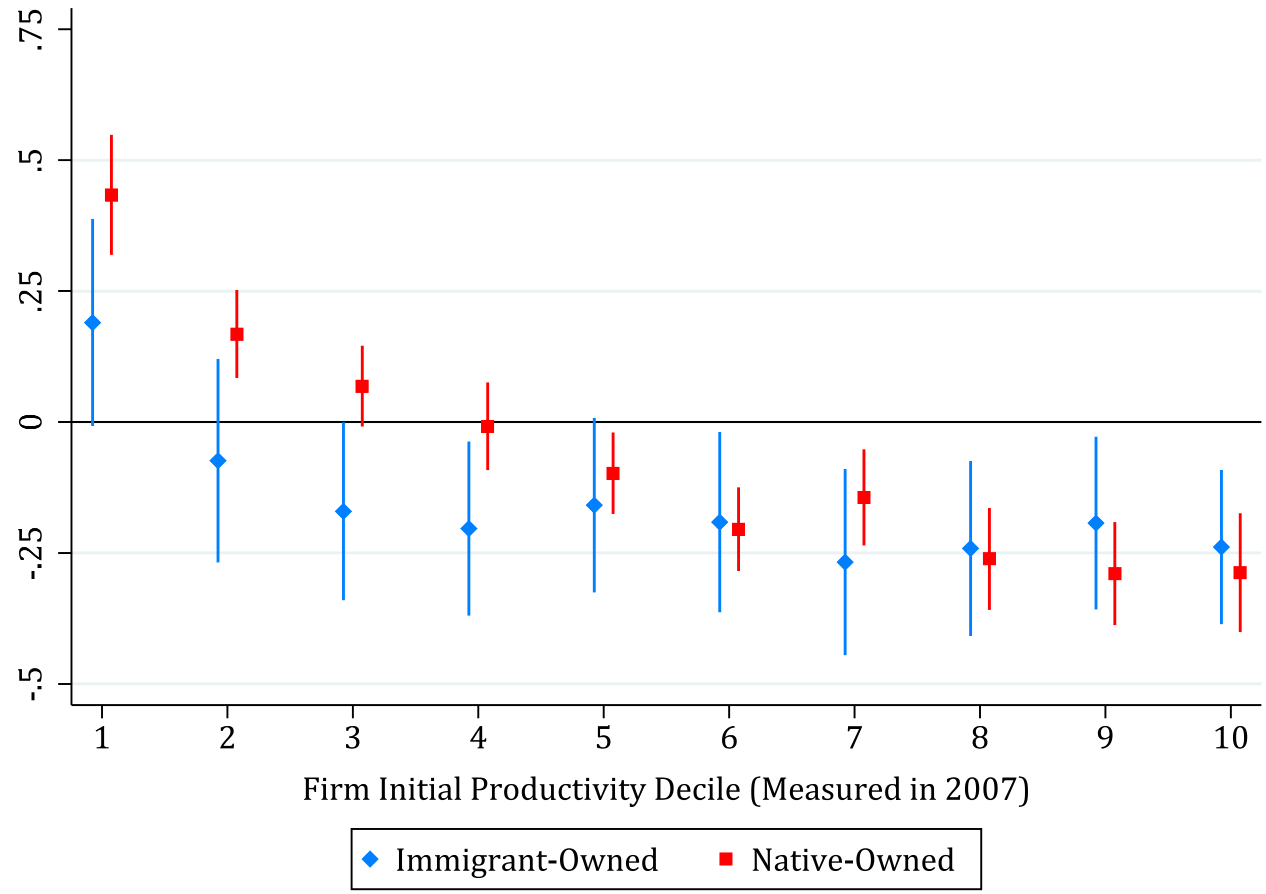
Notes: See Equation (2.1) for specification. Standard errors, clustered at local industry level, in parentheses. Each specification contains 88,806 observations, which represent 722 CZ × 41 industries = 29,602 local industries observed for three decades. In Panel B, the 1st Stage F Statistic for all specifications is 111.4. In all specifications, observations are weighted by 1980 local industry workforce size. All specifications also include control variables for start-of-period college share, start-of-period self-employment share, and start-of-period under-40 share in the local industry, along with a Bartik labor demand control and shift-share controls for exposure to international trade. Finally, all specifications also include commuting-zone-by-year and region-by-industry-by-year fixed effects. IV specifications also include a control variable that interacts the sum of instrument shares with year fixed effects. Rounding errors account for slight discrepancies between (2) - (3) and (1). * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table A4: Decomposition of Immigrant-Induced Job Creation (Figure 5)

	DHS Growth Rate in Employment		Contribution of:						
			Entrants		Exits		Continuers		Residual
	(1)	=	(2)	−	(3)	+	(4)	+	(5)
Panel A: OLS									
ΔI_{gkt}: Immigrant Inflows per Initial Worker	0.6615***		0.2249***		-0.2982***		0.1331***		0.0053
	(0.0592)		(0.0342)		(0.0290)		(0.0260)		(0.0039)
Percent of Total	100		34.00		45.08		20.12		0.01
Panel B: Emigrants IV									
ΔI_{gkt}: Immigrant Inflows per Initial Worker	0.4330***		0.0720		-0.3515***		0.0272		-0.0177
	(0.1279)		(0.0839)		(0.0727)		(0.0708)		(0.0140)
Percent of Total	100		16.63		81.18		6.28		-3.92

Notes: See Equation (2.1) for specification. Standard errors, clustered at local industry level, in parentheses. Each specification contains 88,806 observations, which represent 722 CZ × 41 industries = 29,602 local industries observed for three decades. In Panel B, the 1st Stage F Statistic for all specifications is 111.4. In all specifications, observations are weighted by 1980 local industry workforce size. All specifications also include control variables for start-of-period college share, start-of-period self-employment share, and start-of-period under-40 share in the local industry, along with a Bartik labor demand control and shift-share controls for exposure to international trade. Finally, all specifications also include commuting-zone-by-year and region-by-industry-by-year fixed effects. IV specifications also include a control variable that interacts the sum of instrument shares with year fixed effects. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Figure A1: Effect of Immigration on Establishment Exit, Stratifed by Initial Firm Productivity and Ownership Nativity (OLS Results Corresponding to Figure 9, 2007–2017)



Notes: See Equation (5.11) for specification. Each coefficient represents the effect (in percent) of a 1% immigration shock to a local industry's workforce on the probability that a given establishment is no longer operating in a given year. An establishment is classified as "operating" if it has positive payroll and employment. Establishments are split into deciles based on their parent firm's national rank in sales per worker within 5-digit NAICS code and age bin in 2007. Ownership nativity also defined at the firm level in 2007. Spikes indicate 95% confidence intervals (standard errors clustered at the local industry level). Specification covers 773,000 establishments (rounded to avoid disclosure concerns), weighted by 2007 SBO survey weight. Each specification includes control variables for 2007 college share, 2007 self-employment share, and 2007 under-40 share in the local industry, along with a Bartik labor demand control and shift-share controls for exposure to international trade. Finally, each specification also includes establishment fixed effects, commuting-zone-by-year fixed effects, and region-by-industry-by-year fixed effects. See Figure 9 for corresponding results estimated using IV.

Table A5: The Effect of Immigration on Establishment Exit—Robustness (Corresponding OLS Results to Table 4, 2000–2015)

	Outcome: $\mathbb{1}\{\text{Not Operating}\}_{et}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
I_{gkt}: Immigrant Workers per 2000 Workforce ($\hat{\eta}_0$)	0.4340***	0.4167***	0.6360***	0.4217***	0.5752***	0.5074***	0.7180***
	(0.0409)	(0.0414)	(0.0519)	(0.0243)	(0.0404)	(0.0246)	(0.0351)
$I_{gkt} \times \left[\text{Prod. Pctl./100}\right]_e$ ($\hat{\eta}_1$)	-1.853***	-1.750***	-2.206***	-1.358***	-2.050***	-1.100***	-2.695***
	(0.1330)	(0.1366)	(0.1594)	(0.088)	(0.1254)	(0.0878)	(0.1336)
$I_{gkt} \times \left[\text{Prod. Pctl./100}\right]_e^2$ ($\hat{\eta}_2$)	1.308***	1.212***	1.178***	0.612***	1.147***	-0.0453	1.707***
	(0.1056)	(0.1085)	(0.1270)	(0.0755)	(0.0964)	(0.0778)	(0.1053)
N_{gkt}: Native Workers per 2000 Workforce		0.0150***					
		(0.0034)					
$N_{gkt} \times \left[\text{Prod. Pctl./100}\right]_e$		-0.1131***					
		(0.0159)					
$N_{gkt} \times \left[\text{Prod. Pctl./100}\right]_e^2$		0.1056***					
		(0.0138)					
Implied Crossing Point Pctl.	29.6	30.1	35.6	37.3	34.9	45.3	33.9
Productivity Measure	Revenues p.w.	Revenues p.w.	Revenues	Employment	Payroll p.w.	Employment	Payroll p.w.
Productivity Measurement Level	Firm	Firm	Firm	Firm	Firm	Estab.	Estab.
Productivity Rank Bin	Age Group × NAICS-5D	Age Group × NAICS-5D	Age Group × NAICS-5D	Age Group × NAICS-5D	Age Group × NAICS-5D	Local Industry	Local Industry
Establishments	4,739,000	4,739,000	4,739,000	6,180,000	6,180,000	6,180,000	6,180,000

Notes: See Equation (4.3) for specification. Standard errors, clustered at local industry level, in parentheses. Each establishment followed for four observations ($t \in \{2000, 2005, 2010, 2015\}$). Establishment and observation counts rounded to avoid disclosure concerns. "Implied Crossing Point Pctl." refers to the rank percentile of productivity that the effect of I_{gkt} on $\mathbb{1}\{\text{Not Operating}\}_{et}$ turns from positive to negative. Percentiles (Pctl.) are determined based on a unit's rank of "Productivity Measure" within "Productivity Rank Bin," where a unit corresponds to the "Productivity Measurement Level" row. "p.w." stands for per worker. All specifications also include control variables for 2000 college share, 2000 self-employment share, and 2000 under-40 share in the local industry, along with a Bartik labor demand control and shift-share controls for exposure to international trade. All specifications also include establishment fixed effects, commuting-zone-by-year fixed effects, and region-by-industry-by-year fixed effects. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

B.2 The Effect of Immigrant Worker Inflows on Establishment Presence: Heterogeneity

Figure A2 displays the results of additional heterogeneity analyses across decade (t), industry group (k), and geographic region ($r(g)$). The plotted coefficients come from estimating Equation (2.1) with

$$\Delta y_{gkt} = \frac{\Delta \text{Estabs}_{gkt}}{\text{Estabs Denom}_{gkt}}$$

as the outcome, but conditional on g, k, or t belonging to a given group.[57]

The first primary takeaway from Figure A2 is that effects are generally not driven by one group. The South Region is the only sub-group for which we cannot reject the null hypotheses of no effect of immigrant worker inflows on increased establishment presence. There is a broad-based effect of immigrant worker inflows on establishment presence.

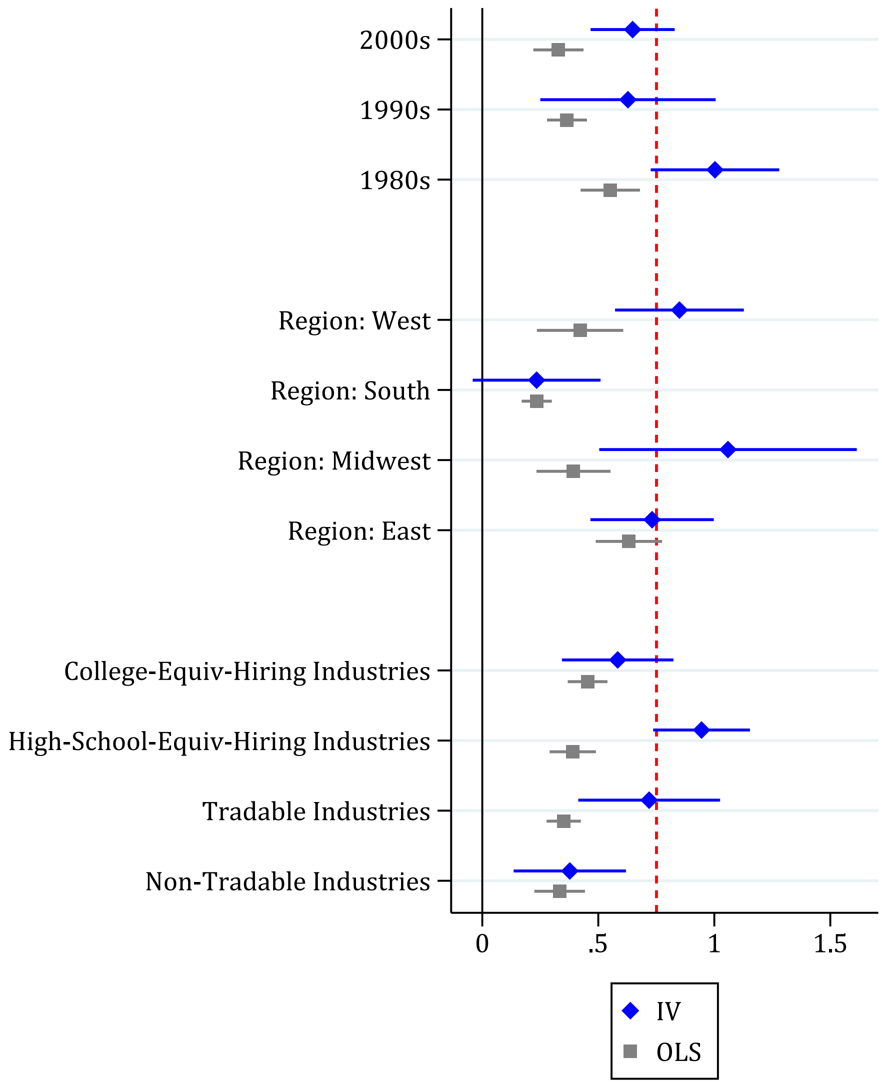
More nuanced takeaways arise from differences across groupings. Of particular interest are the industry heterogeneities presented at the bottom of Figure A2. Going from the bottom up, the first set of analyses designates each industry group (Table A1) as tradable or non-tradable, then estimates Equation (2.1) separately for each set of industry groups. I generate this designation by aggregating 1980 traded and non-traded employment within each industry group based on the Porter classification system for 6-digit NAICS codes (Porter, 2003). Each industry group is then designated as tradable if more than 50 percent of its employment was in a tradable 6-digit NAICS code in 1980, and vice versa. The result of this exercise finds a larger effect in tradable industry groups, but a large effect for non-tradable industry groups as well. The larger effect in the tradable sector comports with findings in both Olney (2013) and Burstein et al. (2020), but the effect for non-tradable industry groups contrasts Olney (2013).

I also designate industry groups based on whether they tend to hire higher- or lower-educated workers. Similar to Doms et al. (2010), I do this by assigning industry groups with below the median share (across industry groups) of college equivalent[58] workers in 1980 the "high-school-equivalent-hiring" designation and industry groups with above the median share the "college-equivalent-hiring" designation. This exercise shows a larger effect in high-school-equivalent hiring industry groups. While this results may be surprising at first blush, it comports with Figure 2, which shows that immigrant worker inflows pushed in by $\Delta z_{gkt}^{\text{Emigrants}}$ are tilted towards high-school-equivalent workers. To the extent that immigrant workers are being absorbed on the extensive margin—as found in Section (3.2)—we would then expand there to be more required expansion in these industries.

[57] E.g., if $r(g)$ = West.

[58] 0.5 times the number of workers with "Some College" plus all workers with at least a four-year college degree.

Figure A2: Effect of Immigration on Establishment Presence in Local Industries—Heterogeneity (1980–2010)

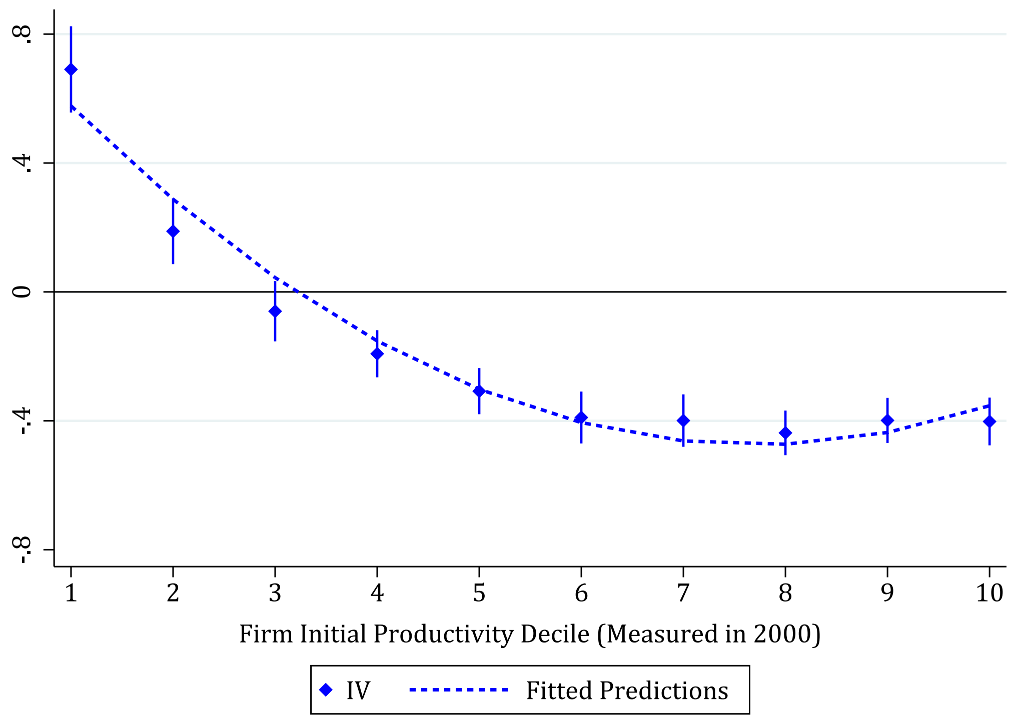


Notes: See Equation (2.1) for specification. Each coefficient represents the effect of a 1% immigration shock to a local industry's workforce on the DHS growth rate in the number of operating establishments in that local industry among observations that fall into the specified group. Spikes around plotted coefficients represent 95% confidence intervals (standard errors clustered at the local industry level). Dashed red reference line is the estimated effect among all observations (Column 3 from Table 1).

B.3 Immigrant Workers and Establishment Exit: Model Fit

In Figure A3, I fit the results from Column 1 of Table 4 at the middle percentile of each decile (e.g., 5th percentile for the 1st decile, 15th percentile for the 2nd decile, etc.) to the IV results from Figure 6. It demonstrates that Equation 4.3 is able capture the less parametric results generated by Equation 4.2 well.

Figure A3: Effect of Immigration on Stratified Establishment Exit—Model Fit of Equation (4.3) Results to Equation (4.2) Results (IV Results, 2000-2015)



Notes: See Equations (4.2) and (4.3) for specification. Each plotted coefficient is estimated from Equation (4.2) and represents the effect (in percent) of a 1% immigration shock to a local industry's workforce on the probability that a given establishment is no longer operating. An establishment is classified as "operating" if it has positive payroll and employment. Establishments are split into deciles based on their parent firm's national rank in revenues per worker within 5-digit NAICS code and age bin in 2000. Spikes indicate 95% confidence intervals (standard errors clustered at the local industry level). Fitted predictions are estimated from Equation (4.3). Each specification covers 4,739,000 establishments (rounded to avoid disclosure concerns), followed every five years until 2015. The 1st Stage F Statistic for the specification that generates coefficient estimates from Equation (4.2) is 8.91, and the 1st Stage F statistic for the specification that generates fitted predictions is 29.72. Each specification is estimated using 88,806 observations that represent 722 CZ $\times$ 41 industries = 29,602 local industries observed for three decades, weighted by 1980 workforce size in the local industry. Each specification includes control variables for 2000 college share, 2000 self-employment share, and 2000 under-40 share in the local industry, along with a Bartik labor demand control and shift-share controls for exposure to international trade. Each specification also includes establishment fixed effects, commuting-zone-by-year fixed effects, and region-by-industry-by-year fixed effects.

B.4 Immigrant Workers and Establishment Exit: Heterogeneity

Table A6: The Effect of Immigration on Stratified Establishment Exit—Heterogeneity (2000–2015)

	Outcome: $\mathbb{1}\{\text{Not Operating}\}_{et}$							
Group:	High-School-Equiv. Hiring	College-Equiv. Hiring	Non-tradable	Tradable	East	Midwest	South	West
Panel A: OLS								
I_{gkt}: Immigrant Workers per 2000 Workforce ($\hat{\eta}^0$)	0.4257***	0.4472***	0.4550***	0.3487***	0.5381***	0.6607***	0.4587***	0.3379***
	(0.0640)	(0.0227)	(0.0546)	(0.0463)	(0.0776)	(0.0546)	(0.0227)	(0.0790)
$I_{gkt} \times \left[\text{Prod. Pctl.}/100\right]_e$ ($\hat{\eta}^1$)	-1.780***	-1.964***	-1.965***	-1.421***	-1.937***	-2.749***	-2.038***	-1.486***
	(0.2066)	(0.0836)	(0.1628)	(0.1543)	(0.2765)	(0.2110)	(0.1001)	(0.2518)
$I_{gkt} \times \left[\text{Prod. Pctl.}/100\right]_e^2$ ($\hat{\eta}^2$)	1.239***	1.408***	1.389***	1.009***	1.229***	1.921***	1.538***	1.028***
	(0.1641)	(0.0653)	(0.1302)	(0.1064)	(0.1992)	(0.1677)	(0.0828)	(0.1999)
Implied Crossing Point Pctl.	30.3	28.7	29.2	31.6	36.0	30.5	28.7	28.3
Panel B: IV								
I_{gkt}: Immigrant Workers per 2000 Workforce ($\hat{\eta}^0$)	0.8193***	0.7244***	0.5893***	1.897***	0.6826***	1.456***	0.7380***	0.5582***
	(0.0962)	(0.0851)	(0.0793)	(0.1849)	(0.1185)	(0.4934)	(0.0813)	(0.0882)
$I_{gkt} \times \left[\text{Prod. Pctl.}/100\right]_e$ ($\hat{\eta}^1$)	-3.179***	-3.490***	-2.923***	-7.220***	-2.854***	-7.045***	-3.194***	-2.771***
	(0.2966)	(0.3015)	(0.1902)	(0.6550)	(0.3744)	(1.427)	(0.2686)	(0.2047)
$I_{gkt} \times \left[\text{Prod. Pctl.}/100\right]_e^2$ ($\hat{\eta}^2$)	2.081***	2.492***	2.031***	4.923***	1.795***	5.125***	2.395***	1.958***
	(0.2209)	(0.2241)	(0.1439)	(0.5260)	(0.2536)	(1.030)	(0.2148)	(0.1454)
Implied Crossing Point Pctl.	32.8	25.3	24.2	34.3	29.3	25.3	29.7	24.3
1st Stage F Statistic	7.55	20.08	29.11	4.21	19.3	7.64	4.45	11.5

Notes: See Equation (4.3) for specification. Standard errors, clustered at local industry level, in parentheses. Each establishment followed for four observations ($t \in \{2000, 2005, 2010, 2015\}$). Each specification estimated using 4,739,000 establishments. "Implied Crossing Point Pctl." refers to the rank percentile of productivity that the effect of I_{gkt} on $\mathbb{1}\{\text{Not Operating}\}_{et}$ turns from positive to negative. Productivity percentiles measured based on parent firm's revenue per worker rank within 5-digit NAICS code by age group bin. All specifications include control variables for 2000 college share, 2000 self-employment share, and 2000 under-40 share in the local industry, along with a Bartik labor demand control and shift-share controls for exposure to international trade. Each specification also includes establishment fixed effects, commuting-zone-by-year fixed effects, and region-by-industry-by-year fixed effects. Vector of instrumental variables in Panel Bis $\left(z_{gkt}^{\text{Emigrants}}, z_{gkt}^{\text{Emigrants}} \times \left[\text{Prod. Pctl.}/100\right]_e, z_{gkt}^{\text{Emigrants}} \times \left[\text{Prod. Pctl.}/100\right]_e^2\right)$. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

B.5 Effect of Immigration on Establishment Exit in the SBO Sub-Sample

Here, I present results of estimating Equation (4.1) and Equation (4.3) using the sample of establishments from the representative sample of non-public, operating firms in 2007 contained in the SBO. These analyses probe whether the effect of immigrant worker inflows on establishment exit changes when:

- The study period changes from 2000–2015 to 2007–2017
- I use the representative, 2007 SBO sample instead of 1) the full-count dataset studied in Section 4.1 (Table 3, Column 1); or 2) the potentially non-representative subset of the full-count dataset that can be linked to revenue information in 2000 studied in Section 4.2.3.
- I use $z_{gk,t-2}^{\text{Emigrants}}$ as the instrumental variable instead of $z_{gkt}^{\text{Emigrants}}$

Table A7 presents the results. They indicate that the effects found in Sections 4.1, 4.2.3, and 5.5.2 are all robust to these changes: immigrant worker inflows have similar impacts on the probability of operation and similar form of stratification by productivity ranks across samples.

Table A7: The Effect of Immigration on Establishment Exit—SBO Sub-Sample (2007–2017)

Outcome:	I_{gkt}	Outcome: $\mathbb{1}\{\text{Not Operating}\}_{et}$			
	(1)	(2)	(3)	(4)	(5)
$z_{gk,t-2}$: Lagged Emigrants Instrument	0.1084***				
	(0.0084)				
I_{gkt}: Immigrant Workers per 2007 Workforce		-0.0440***	-0.2298***	0.4417***	2.955***
		(0.0115)	(0.0720)	(0.0564)	(0.3611)
$I_{gkt} \times \left[{}^{\text{Prod. Pctl.}}/{}_{100}\right]_e$				-1.764***	-11.48***
				(0.2254)	(1.236)
$I_{gkt} \times \left[{}^{\text{Prod. Pctl.}}/{}_{100}\right]_e^2$				1.103***	7.845***
				(0.1904)	(0.8853)
Implied Crossing Point Pctl.	—	—	—	31.1	33.3
Instrument	N/A: 1st Stage	N/A: OLS	$z^{\text{Emigrants}}_{gk,t-2}$	N/A: OLS	$z^{\text{Emigrants}}_{gk,t-2}$
1st Stage F Statistic	—	—	167.8	—	54.46
Productivity Measure	—	—	—	Sales p.w.	Sales p.w.
Productivity Measurement Level	—	—	—	Firm	Firm
Productivity Rank Bin	—	—	—	Age Group × NAICS-5D	Age Group × NAICS-5D
Establishments	773,000	773,000	773,000	773,000	773,000

Notes: See Equation (4.3) for specification. Standard errors, clustered at local industry level, in parentheses. Each establishment followed for three observations ($t \in \{2007, 2012, 2017\}$), weighted by 2007 SBO survey weight. Establishment and observation counts rounded to avoid disclosure concerns. "Implied Crossing Point Pctl." refers to the rank percentile of productivity that the effect of I_{gkt} on $\mathbb{1}\{\text{Not Operating}\}_{et}$ turns from positive to negative. Percentiles (Pctl.) are determined based on a unit's rank of "Productivity Measure" within "Productivity Rank Bin," where a unit corresponds to the "Productivity Measurement Level" row. "p.w." stands for per worker. All specifications also include control variables for 2007 college share, 2007 self-employment share, and 2007 under-40 share in the local industry, along with a Bartik labor demand control and shift-share controls for exposure to international trade. All specifications also include establishment fixed effects, commuting-zone-by-year fixed effects, and region-by-industry-by-year fixed effects. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

B.6 Instrument Vetting

This section presents several analyses and visualizations that bolster the case for $\Delta z_{gkt}^{\text{Emigrants}}$ as a relevant and valid instrument. Where indicated, some of these analyses are conducted using publicly-available data in order to avoid excessive disclosure burden on the Census Bureau.

B.6.1 Confounding Short- and Long-Run Effects (Jaeger et al., 2018)

Jaeger et al. (2018) broach a concern that arises from serial correlation in the "shift" component of "shift-share" instruments—one that is particularly concerning with regards to the standard immigration shift-share instrument. When this shift component is excessively serially correlated over time, estimated parameters like β in in Equation (2.1) can confound short- and long-run responses to immigrant inflows. Though this concern is particularly deleterious when wages are the primary outcome variable of interest, it merits consideration in any setting where prior shocks may affect current outcomes.

Jaeger et al. (2018) propose a data-demanding procedure to both test for and account for such concerns, which is to include both the independent variable and its lag, and to instrument for both (i.e., here, include ΔI_{gkt} and $\Delta I_{gk,t-10}$ and use both $\Delta z_{gkt}^{\text{Emigrants}}$ and $\Delta z_{gk,t-10}^{\text{Emigrants}}$ as instruments.

$$\Delta y_{gkt} = \beta_0 \left(\Delta I_{gkt}\right) + \beta_{-1} \left(\Delta I_{gk,t-10}\right) + \Gamma X_{gkt} + \alpha_{gt} + \alpha_{r(g),k,t} + \varepsilon_{gkt} \tag{B.1}$$

Table A8 display comparable estimates for the two decades covered by 1990–2010, and with Column 6 implementing the double-instrumentation strategy proposed by in Jaeger et al. (2018) to account for serial correlation that can undermine shift-share-based estimate interpretation. The relative strength of the 1st Stage F statistic in Column 6 is reassuring, as more standard shift-share instruments often collapse in strength under this procedure. In line with Jaeger et al. (2018), I hypothesize that the reason is that emigrant outflows to non-U.S. locations are less serially correlated across time across origin countries o than are immigrant inflows to the U.S., which are dominated by Mexico. The double-instrumentation effect estimate in Column 6 is less precise and attenuated by roughly one-fourth relative to the estimate in Column 5, but the estimate in Column 5 is well within the confidence band. Because this procedure forces me to drop the 1980s as part of the study period, and because I cannot detect strong evidence of bias, I implement single instrumentation throughout the main text of the paper.

Table A8: The Effect of Immigration on Establishment Presence (1990–2010)

Outcome:	DHS Growth Rate in Establishment Count	
	(1)	(2)
ΔI_{gkt}: Immigrant Inflows per Initial Worker	0.6628***	0.5055***
	(0.0873)	(0.1864)
$\Delta I_{gk,t-10}$: Lagged Immigrant Inflows per Initial Worker		0.2333
		(0.2404)
Instrument(s)	$\Delta z_{gkt}^{\text{Emigrants}}$	$\begin{pmatrix} \Delta z_{gkt}^{\text{Emigrants}} \\ \Delta z_{gk,t-10}^{\text{Emigrants}} \end{pmatrix}$
1st Stage F Statistic	90.65	25.09
Within R^2	0.0025	0.0042
Commuting Zone × Year FE	✓	✓
Region × Industry Group × Year FE	✓	✓
Controls	✓	✓
Observations	59,204	59,204

Notes: See Equation (B.1) for specification. Standard errors, clustered at local industry level, in parentheses. Each decade studied contains 722 CZ × 41 industries = 29,602 local industries. In all specifications, observations are weighted by 1980 local industry workforce size. Where indicated, specifications also include control variables for start-of-period college share, start-of-period self-employment share, and start-of-period under-40 share in the local industry, along with a Bartik labor demand control and shift-share controls for exposure to international trade. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

B.6.2 *Correlated Shocks Across gk with Similar Shares (Adao et al., 2019)*

Adao et al. (2019) find that regression residuals can be substantially correlated across areas with similar "share" components in shift-share instruments, invalidating standard inference procedures. Importing their concerns to the current study, any industry-country level shocks that affect outcomes through the presence of base year shares $\pi_{go,1980}$, even if not related to immigration itself, can generate correlated outcomes across areas with similar $\pi_{go,1980}$.

A simple example that could apply here would be a sector-specific trade shock in a given origin country. For example, if Syria experiences a positive trade shock that is independent of Syrian emigration forces, this can affect firm presence in areas heavily populated by Syrians in the U.S.—e.g., Detroit and Boston—through trade linkages, but it is unlikely to have any affect on firm presence in areas like Atlanta or Miami with low Syrian populations. These correlated shocks would not create a bias in β, but would require a modification of standard errors beyond clustering at the commuting zone-industry level, since Detroit and Boston are not even in the same region.

Adao et al. (2019) illustrate this issue by conducting placebo tests in which they replace the "shift" component with white noise and assessing the resulting false rejection rate after multiple simulations of the reduced form regression model. Here, the analogous placebo exercise uses instruments of the form

$$z_{gkt}^{\text{Placebo}} \frac{1}{E_{g,1980}} \sum_{o} \pi_{go,1980} \times \omega_{okt}$$

where ω_{okt} is a random draw from a normal distribution . Each placebo instrument is then used in the reduced form estimating equation:

$$\Delta y_{gkt} = \alpha + \beta^{\text{Placebo}} \left[\Delta z_{gkt}^{\text{Placebo}} \right] + \Gamma X_{gkt} + \alpha_{gt} + \alpha_{r(g)kt} + u_{gkt}$$

with the DHS growth rate in the local industry establishment count and standard errors clustered at the local industry level.

Results from 1,000 placebo simulations are summarized in Table A9. There does not appear to be any evidence for over-rejection from these simulations. I hypothesize that the rich fixed effects structure $\alpha_{gt} + \alpha_{r(g)kt}$ removes much of the correlation across geographies found in Adao et al. (2019) when they study the standard immigration instrument at the gt (rather than gkt) level

Table A9: Summary Characteristics of 1,000 t-Statistics from Adao et al. (2019) Placebo Simulations

	Simulation Result	Target
Mean	-0.0043	0
Standard Deviation	1.007	1
False Rejection Rate: 90% Confidence Interval	0.096	0.1
False Rejection Rate: 95% Confidence Interval	0.040	0.05
False Rejection Rate: 99% Confidence Interval	0.009	0.01

B.6.3 Emigrants IV: Example (Publicly-Available Data)

A simple and relevant example that helps illustrate the utility of $\Delta z_{gkt}^{\text{Emigrants}}$ comes from the housing bubble that crested between 2000 and 2005, largely in the South and West of the U.S. The housing bubble created a large labor demand shock for construction workers in the South and West Census Regions of the U.S., and induced immigrant workers from Mexico to fill this demand—the kind of inflow an instrumental variable should not use for identification of β in Equation (2.1) because it will confound a labor-demand induced immigration with the response of labor demand to immigration. As seen in Panel A of Figure A4, Mexican inflows into the construction sector between 2000 and 2005 were more than 10 times larger than those from the next closest country. As seen in Panel B of Figure A4, general immigrant inflows into the construction sector across commuting zones predominantly took place in housing bubble cities throughout the South and West regions of the country.[59] On the other hand, there is no reason to believe that the U.S. housing bubble would cause large outflows of Mexican emigrants to non-U.S. OECD countries. Furthermore, relative to the national trend, the change in propensity of Mexican immigrants to locate in the construction sector, *outside* of the South and West regions was not unusually strong. Thus, the aggregate component of $\Delta z_{gkt}^{\text{Emigrants}}$ between 2000 and 2005

$$\left[\rho_{ok,t=2005,-r(g)} \times \mathcal{M}_{o,t=2005}^{\text{non-US}} - \rho_{ok,t=2000,-r(g)} \times \mathcal{M}_{o,t=2000}^{\text{non-US}}\right]$$

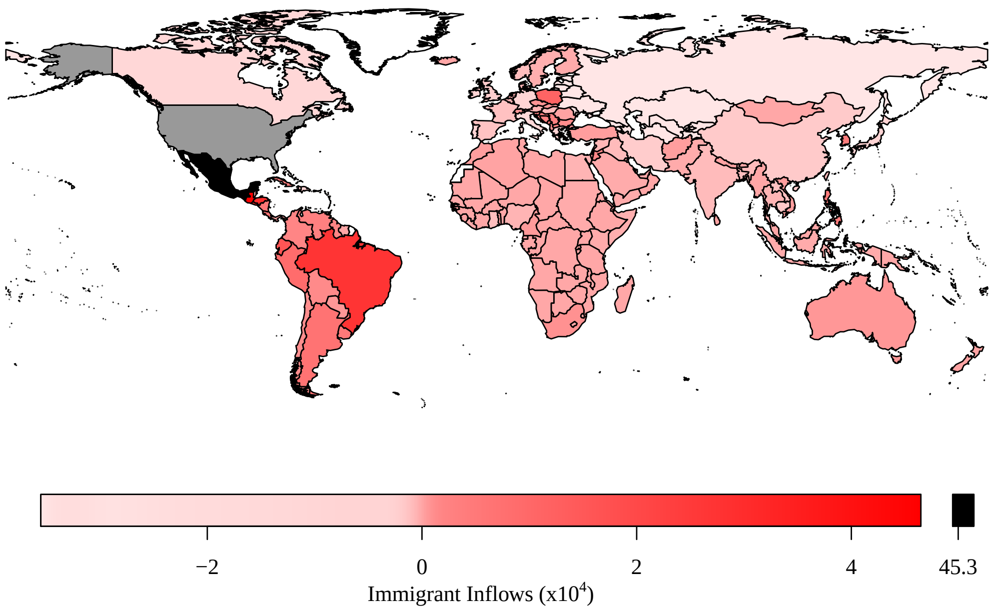
should not reflect the labor demand shocks in construction that were occurring in the South and West of the country at that time. These two factors are illustrated in Panel A of Figure A5, where Mexico has a much more modest aggregate component for the construction sector between 2000 and 2005. The ultimate result of these corrections can be seen in Panel B, where the instrument-predicted immigrant inflows are far less concentrated in bubble cities.

[59] Note that initial shares of Mexican workers, $\pi_{\text{Mexico},g,1980}$, are also high in South and West commuting zones. So, beyond using OLS to estimate Equation (2.1) during this time period, using $\Delta z_{gkt}^{\text{Standard}}$ as an instrument for IV estimation also poses an issue.

Figure A4: Endogenous Immigrant Inflows into the Construction Industry (2000–2005)

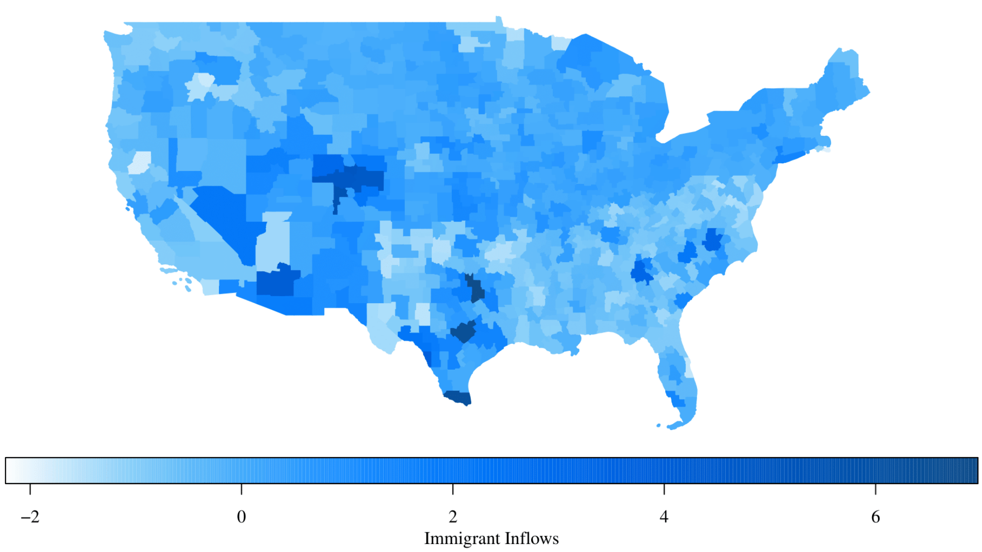
Panel A: Net Immigration by Country, k=Construction

$$\left[I_{ok,2005} - I_{ok,2000}\right]$$



Panel B: Immigrant Inflows into k=Construction

$$\left[I_{gk,2005} - I_{gk,2000}\right]$$

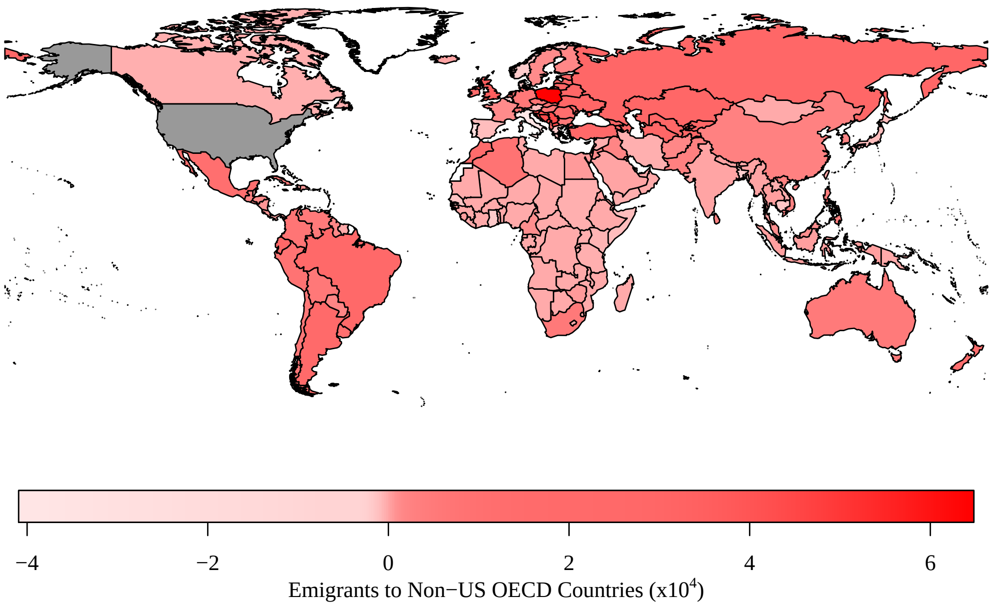


Notes: Net immigration in Panel B is standardized across commuting zones.

Figure A5: Exogenous Immigrant Inflows into the Construction Industry (2000–2005)

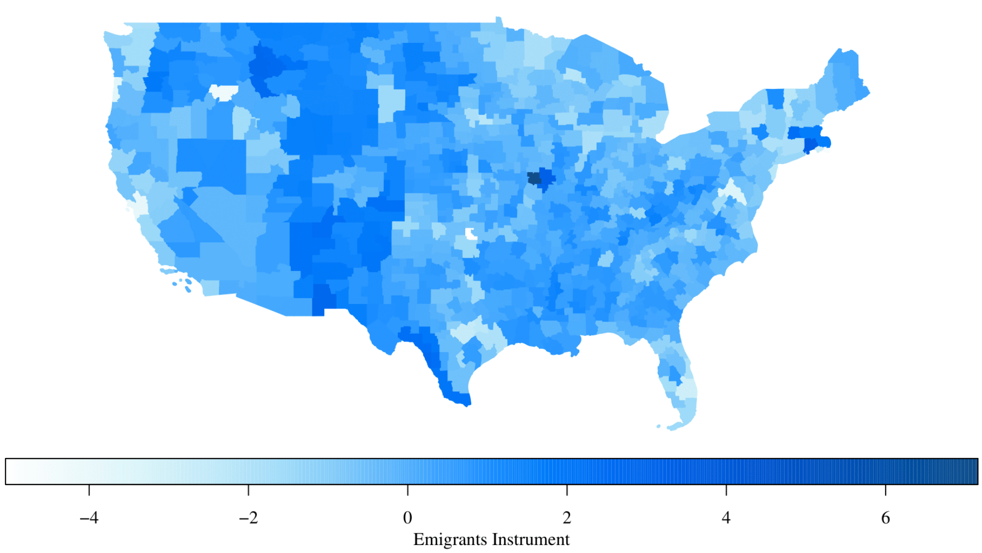
Panel A: Emigration Shock, k=Construction

$$\left[(\rho_{ok,2005} \times M^{\text{non-US}}_{o,2005}) - (\rho_{ok,2000} \times M^{\text{non-US}}_{o,2000})\right]$$



Panel B: Instrument-Predicted Inflows into k=Construction

$$\left[z_{gk,2005} - z_{gk,2000}\right]$$



Notes: Instrument-Predicted Inflows in Panel B are standardized across commuting zones.

B.6.4 *Comparison of* $\Delta z_{gkt}^{Emigrants}$ *with* $\Delta z_{gkt}^{Standard}$ *(Publicly-Available Data)*

In Section 2.2, I argue that the modifications I make to $\Delta z_{gkt}^{\text{Standard}}$ generate a more plausibly exogenous and robust instrument $\Delta z_{gkt}^{\text{Emigrants}}$. I also describe two sources of bias along which we expect improvement: 1) exogeneity of emigration shocks and therefore plausible exogeneity of the instrument itself (Borusyak et al., 2020); 2) less serial correlation in the origin o aggregate component that removes the confounding of short- and long-run responses to immigration (Jaeger et al., 2018).[60] Here, I directly present evidence that $\Delta z_{gkt}^{\text{Emigrants}}$ out-performs $\Delta z_{gkt}^{\text{Standard}}$ along these dimensions. Note that all instruments in these analyses are constructed from publicly-available data from IPUMS-USA (Ruggles et al., 2019).

In order to account for recent best-practice, I construct the more modern version of $\Delta z_{gkt}^{\text{Standard}}$:

$$\Delta z_{gkt}^{\text{Standard}} \equiv \frac{1}{E_{gk,1980}} \sum_{o} \pi_{og,1980} \times \Delta \mathcal{I}_{okt,(-g)}$$

where the aggregate component is now immigrant inflows into industry k from origin o between $t-10$ and t into all commuting zones other than g. This is a common practice that aims to eliminate commuting zone g factors from influencing inflows. As seen below, however, this modification does not do enough to generate a plausibly exogenous instrument.

The first comparison I conduct is on instrument balance. Because other outcomes are not available publicly-in the pre-period at the local industry level, I compare the performance of both instruments in Equation (2.4) using Census-measured DHS growth in the 1970s (also from IPUMS-USA):

$$\Delta y_{gk,pt}^{\text{Std.}} = \beta_{\text{Balance}} \left(\Delta z_{gkt}\right) + \alpha_{gt} + \alpha_{r(g),kt} + \varepsilon_{gkt}$$

As in Figure 1, I also compare the effect of each instrument on a "true" outcome – the standardized version of the DHS growth rate in employment during each study period decade.

Figure A6 presents the results, comparing instrument balance between $\Delta z_{gkt}^{\text{Standard}}$ and $\Delta z_{gkt}^{\text{Emigrants}}$ using publicly-available data. It finds that there is substantially more negative pre-period employment growth among local industries with higher values of $\Delta z_{gkt}^{\text{Standard}}$, and—as in Figure 1—no evidence of such a pre-trend for $\Delta z_{gkt}^{\text{Emigrants}}$. That the correlation between $\Delta z_{gkt}^{\text{Standard}}$ and pre-period employment growth is negative may, at first blush, be less concerning than a corresponding positive correlation. Nonetheless, the statistically significance does suggest non-randomness in $\Delta z_{gkt}^{\text{Standard}}$ and the sign suggests the possibility that mean reversion may contaminate results estimated using $\Delta z_{gkt}^{\text{Standard}}$.

The second comparison I conduct is on resilience to the double-instrumentation procedure suggested in Jaeger et al. (2018). For this purpose, I estimate Equation (B.1) using the DHS growth rate in County Business Pattern (CBP) employment growth:

$$\Delta y_{gkt} = \beta_0 \left(\Delta I_{gkt}\right) + \beta_{-1} \left(\Delta I_{gk,t-10}\right) + \Gamma X_{gkt} + \alpha_{gt} + \alpha_{r(g),k,t} + \varepsilon_{gkt}$$

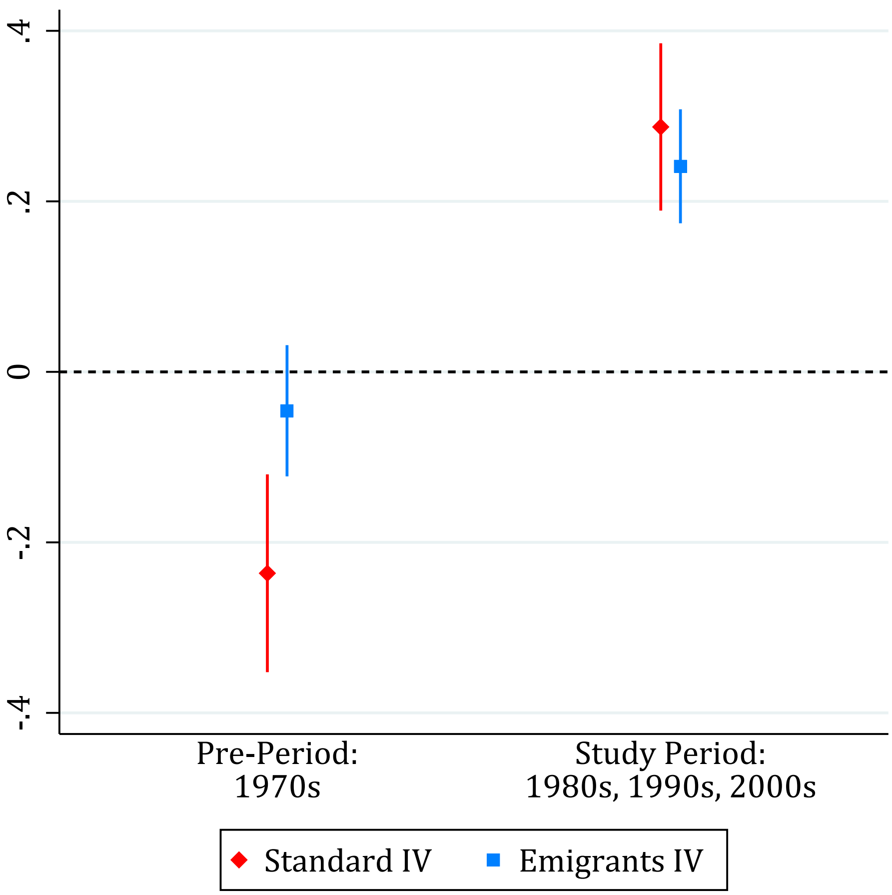
CBP employment is available for consistent industry classifications thanks to crosswalks provided in Eckert et al. (2020), and represents the closest publicly-available analog to employment growth measured from the LBD.

Table A10 presents the results of this exercise for the 1990s and 2000s. Reassuringly, $\Delta z_{gkt}^{\text{Emigrants}}$ delivers similar results in publicly-available data that it does using restricted-access Census data.

[60] Note that both instruments share the same "share" component, so the correlated shocks concern broached by Adao et al. (2019) applies equally and is alleviated by the simulations in Section B.6.2.

This is true both in terms of the impact of immigration on employment growth estimated using single-instrumentation and and the resilience of $\Delta z_{gkt}^{\text{Emigrants}}$ in to double-instrumentation—seen in the First Stage F statistic in Column 3.[61] In line with Jaeger et al. (2018), $\Delta z_{gkt}^{\text{Standard}}$ collapses under double-instrumentation, producing nonsensical results and a non-existent first stage. This stark difference is a strong argument in favor of $\Delta z_{gkt}^{\text{Emigrants}}$.

Figure A6: Balance Tests on Employment Growth—Comparison Across Instruments



Notes: Outcome variable calculated from 1970 and 1980 public-use Census data. Each specification is estimated using 88,806 observations that represent 722 CZ × 41 industries = 29,602 local industries observed for three decades, weighted by 1980 workforce size in the local industry. Each outcome is a standardized version of the indicated variable, regressed on the instrumental variable, $\Delta z_{gkt}^{\text{Emigrants}}$, along with commuting-zone-by-year and region-by-year-by-industry fixed effects. No additional controls are included in these specifications. All variables constructed using data from IPUMS-USA.

[61]It does appear that measurement error may be a larger factor in publicly-available data, producing instrumental variable estimates that are larger than OLS estimates. This is not the case in the restricted_access data (see Figure 5.

Table A10: The Effect of Immigration on Employment (1990–2010)–Comparison Across Instruments

	Outcome: DHS Growth Rate in Employment				
	(1)	(2)	(3)	(4)	(5)
ΔI_{gkt}: Immigrant Inflows per Initial Worker	0.480***	0.552***	0.841***	0.594***	-11.26
	(0.0397)	(0.110)	(0.255)	(0.124)	(204.1)
$\Delta I_{gk,t-10}$: Lagged Immigrant Inflows per Initial Worker			-0.526*		11.96
			(0.280)		(205.0)
Instrument(s)	None—OLS	$\Delta z^{\text{Emigrants}}_{gkt}$	$\begin{pmatrix}\Delta z^{\text{Emigrants}}_{gkt} \\ \Delta z^{\text{Emigrants}}_{gk,t-10}\end{pmatrix}$	$\Delta z^{\text{Standard}}_{gkt}$	$\begin{pmatrix}\Delta z^{\text{Standard}}_{gkt} \\ \Delta z^{\text{Standard}}_{gk,t-10}\end{pmatrix}$
1st Stage F Statistic	—	150.5	23.92	84.48	0.002
Observations	59,204	59,204	59,204	59,204	59,204

Notes: See Equation (B.1) for specification. Standard errors, clustered at local industry level, in parentheses. Each decade studied contains 722 CZ × 41 industries = 29,602 local industries. In all specifications, observations are weighted by 1980 local industry workforce size. Specifications also include control variables for start-of-period college share, start-of-period self-employment share, and start-of-period under-40 share in the local industry, along with a Bartik labor demand control and shift-share controls for exposure to international trade. All specifications also include region-by-industry-by-year fixed effects and commuting-zone-by-year fixed effects. Outcome variable constructed from County Business Pattenns in conjunction with crosswalks provided by Eckert et al. (2020). Independent variables and instrumental variables constructed from IPUMS-USA. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

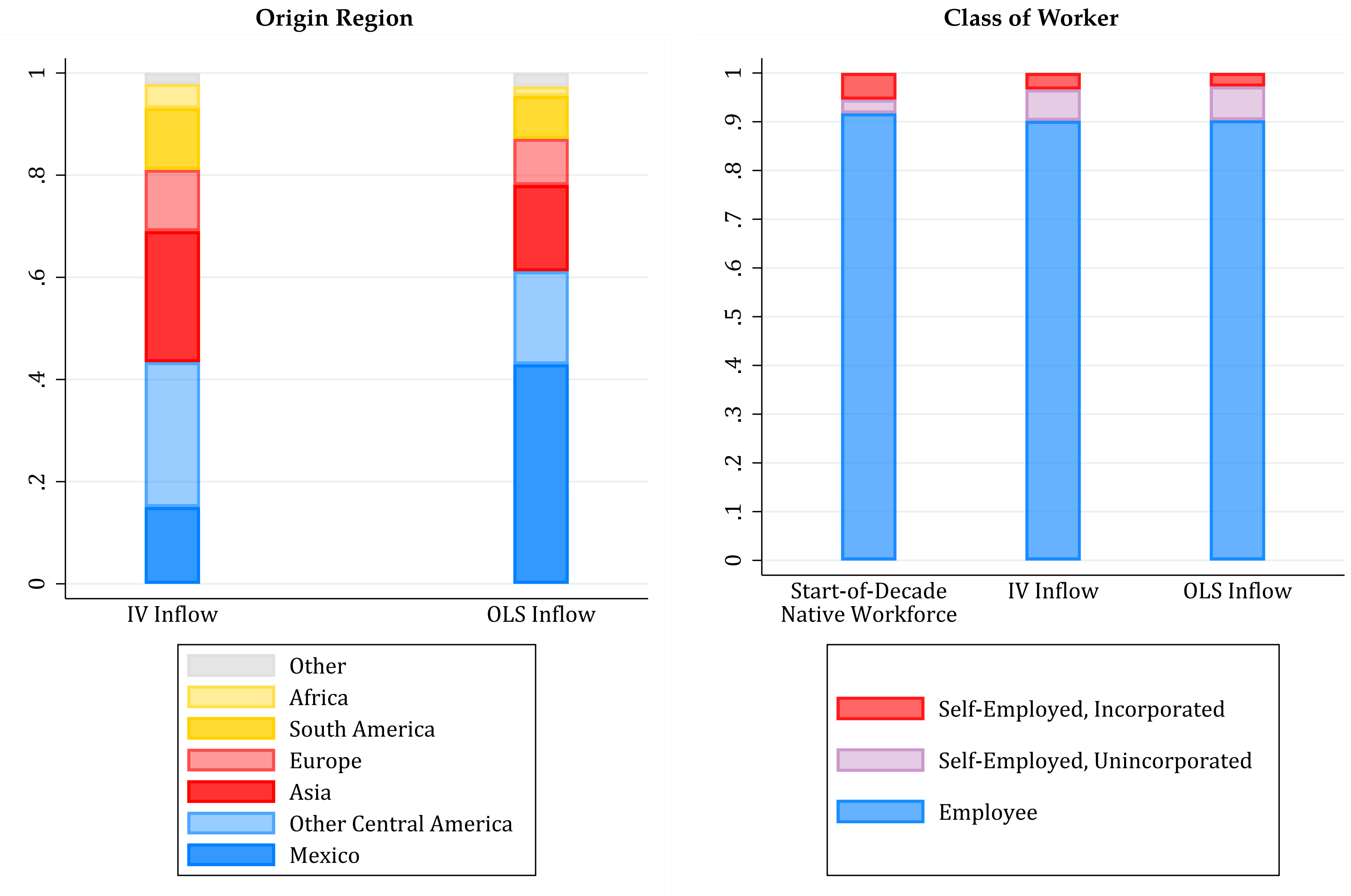
B.6.5 Additional Characteristics of Inflows (Publicly-Available Data)

The following figure decomposes immigrant inflows in ΔI_{gkt} by various characteristics based on the following specification:

$$\Delta I_{gkt}^{\text{Characteristic}} = \alpha + \beta \left[\Delta I_{gkt} \right] + \Gamma X_{gkt} + \alpha_{gt} + \alpha_{d(g)kt} + \varepsilon_{gkt} \quad \text{(B.2)}$$

Equation B.2 exploits the adding-up property of linear regression for mutually exclusive and exhaustive groupings, decomposing how many workers of each characteristic are brought in by each immigrant, on average. Figure 2 of the main text already uses Equation (B.2) to decompose the composition immigrant inflows into educational categories. Note that these regressions are estimated using publicly available data from IPUMS-USA (Ruggles et al., 2019). IV estimates are generated using $\Delta z_{gkt}^{\text{Emigrants}}$ as the instrumental variable.

Figure A7: Additional Characteristics of Immigrant Inflows



B.7 Back-of-the-Envelope Calculations

B.7.1 Immigrant Entrepreneurship and the Effect of Immigrant Worker Inflows on Establishment Entry

I start by noting that results in Appendix Section B.6.5, which show that each immigrant worker pushed in by $\Delta z_{gkt}^{\text{Emigrants}}$ represents 0.903 employees and 0.097 self-employed workers. I then use an estimate of 2 percent for the immigrant entrepreneurship rate (see Figure 3c in Kerr and Kerr, 2016). Kerr and Kerr (2016) also estimate that 68% of immigrant firm owners came from those classifying themselves as "employees" in the 2000 Decennial Census. I plug these numbers in to estimate

$$\begin{aligned}\mathbb{P}[\text{Employee}]\mathbb{P}[\text{Firm Owner}|\text{Employee}] &= \mathbb{P}[\text{Firm Owner}]\mathbb{P}[\text{Employee}|\text{Firm Owner}] \\ &\approx 0.02 \times 0.68 = 0.0136\end{aligned}$$

Kerr and Kerr (2016) also estimate that 29% of immigrant firm owners came from those classifying themselves as "self-employed" in the 2000 Decennial Census. I plug these numbers in to estimate

$$\begin{aligned}\mathbb{P}[\text{Self-Employed}]\mathbb{P}[\text{Firm Owner}|\text{Self-Employed}] &= \mathbb{P}[\text{Firm Owner}]\mathbb{P}[\text{Self-Employed}|\text{Firm Owner}] \\ &\approx 0.02 \times 0.29 = 0.0058\end{aligned}$$

Then, the probability that a given immigrant pushed in by $\Delta z_{gkt}^{\text{Emigrants}}$ is

$$\begin{aligned}\mathbb{P}[\text{Firm Owner}] =& \mathbb{P}[\text{Employee}]\mathbb{P}[\text{Firm Owner}|\text{Employee}] \\ &+ \mathbb{P}[\text{Self-Employed}]\mathbb{P}[\text{Firm Owner}|\text{Self-Employed}] \\ \approx& 0.0140\end{aligned}$$

There can be multiple firm owners per firm and multiple establishments per firm. Kerr and Kerr (2018) report that 50 percent of new, immigrant owned firms have one owner, 36.9 percent have two owners, and the remainder have at least 3 owners in 2007. This implies a lower bound of 1.628 firm owners per firm, on average. Meanwhile, the publicly-available Business Dynamics Statistics reports that there were 1.28 establishments per firm, on average in 2000. Thus an upper bound on the number of new establishments generated by these 0.0194 firm owners is $0.0140 \times \frac{1.28}{1.628} = 0.011$. 0.011 is 26.7 percent of the total number of establishments generated per immigrant worker. Meanwhile, establishment entry as a whole accounts for 43 percent of this total effect (see Table A3). I therefore conclude that new immigrant entrepreneurship can account for up to $\frac{36.9}{43} = 62$ percent of the effect of immigrant worker inflows on establishment entry.

B.7.2 Contextualizing the Effect of Immigrant Workers on Establishment Exit

Panel B of Table 3 indicates that a one percent shock to relative supply due to immigration decreases the probability that an establishment exits by -0.002139 percentage points over a five year horizon for the time period 2000–2015. The Business Dynamics Statistics indicate that the one-year hazard rate establishments in 2000 was 9.191 percent. Taking this as the probability an establishment exits in a given year, we have the overall probability that an establishment has exited over a five year horizon is

$$\mathbb{P}[\text{Not Operating}_{t+5}|\text{Operating}_t] = 1 - \prod_{s=1}^{5}\left(1 - \underbrace{\mathbb{P}[\text{Exit}_{t+s}|\text{Operating}_{t+s-1}]}_{\approx 0.09191}\right)$$
$$\approx 1 - (1 - 0.09191)^5 = 0.4068$$

So, the percent change in exit probability induced by a one percent shock to relative supply due to immigration relative to this general probability of exit is $100 \times \frac{-0.002139}{0.4068} = -0.52\%$. Note that this is likely a lower bound on the magnitude of the effect. Due to selection, once an establishment survives in a given year, its exit hazard in the next year should be lower. Decreasing $\mathbb{P}[\text{Exit}_{t+s}|\text{Alive}_{t+s-1}]$ as s grows would decrease $\mathbb{P}[\text{Not Alive}_{t+5}|\text{Alive}_t]$, and therefore the denominator in the expression above. Note also that these calculations ignore the possibility that an establishment is "not operating" but has not technically exited (i.e., it treats exit as an absorbing state). The separation between non-operation and exit is an empirical rarity.

C Detailed and Expanded Model

C.1 Setup

Individuals are consumer-employees of type $i \in \{I_e, N_e\}$, with I representing foreign-born individuals and N representing native-born individuals and $e \in \{L, S\}$, where L stands for high school degree or *L*ess and S stands for at least *S*ome college. The mass of each labor type in the economy is fixed and employees supply their labor inelastically—the primary comparative static will increase immigrant mass by increasing both I_S and I_L (as an average immigrant inflow into the U.S. does). Entrepreneurs are indexed by the technology j with which they produce, $j \in \{0, 1\}$. $j = 1$ entrepreneurs choose to produce with a technology that is more immigrant-intensive.

C.1.1 Consumer Preferences

Consumer preferences are uniform across consumers. Preferences across firms are of the CES form:

$$\mathcal{U} = \left[F^{\frac{\eta-1}{\mu}} \int_0^F Q(f)^{\frac{\mu-1}{\mu}} df \right]^{\frac{\mu}{\mu-1}}$$

where F is the mass of firms, $Q(f)$ is the amount demanded by consumers at firm f. $\mu > 1$ is the elasticity of substitution of consumption across firms. When $\eta = 1$, consumers have a taste for variety, as the usual CES preference model of monopolistic competition dictates. This taste for variety generates external scale effects through which an increasing market size increases welfare. When $\eta = 0$, we shut down this channel and focus on the firm productivity distribution (see, e.g., Egger and Kreickemeier, 2009).

This results in the following demand curves for each firm, which are downward sloping due to product differentiation and substitutability across goods:

$$Q(f) = YF^{\eta-1}P^{\mu-1}p(f)^{-\mu} \tag{C.1}$$

where $p(f)$ is the price charged by firm f and Y is total consumer spending, and the price index P is given by $P^{1-\mu} \equiv F^{\eta-1} \int_0^F p(f)^{1-\mu} df$.

C.1.2 Firms

Firms have some market power but are non-strategic and take their downward-sloping demand curves as given—a typical monopolistic competition setup. Firm production functions are given by

$$\begin{aligned}
Q_j(z) &= zq_j(z) \\
q_j(z) &= \left[a_L \left(L_j(z)\right)^{\frac{\sigma_E-1}{\sigma_E}} + \left(S_j(z)\right)^{\frac{\sigma_E-1}{\sigma_E}} \right]^{\frac{\sigma_E}{\sigma_E-1}} \\
L_j(z) &= \left[b_j \left(I_{Lj}(z)\right)^{\frac{\sigma_I-1}{\sigma_I}} + \left(N_{Lj}(z)\right)^{\frac{\sigma_I-1}{\sigma_I}} \right]^{\frac{\sigma_I}{\sigma_I-1}} \\
S_j(z) &= I_{Sj}(z) + N_{Sj}(z)
\end{aligned}$$

where z is a draw of total factor productivity, $q_j(z)$ is a CES aggregator of less-educated labor (L_j) and more-educated labor (S_j), and the lower education labor group is itself a CES aggregator of immigrant and native labor. I consider high education workers to be indistinguishable across nativity.[62] The elasticities σ_E and σ_I govern how substitutable workers of different education and different nativities are, respectively. z is drawn from the same Pareto distribution with shape parameter ϕ and minimum value m. This draw endogenously determines whether or not the entrepreneur produces and with which technology. The parameter a_L governs relative productivity across less- and more-educated workers.

The key difference across firms producing with different technologies is the parameter b_j within the low-education aggregator. I assume $j = 1$ firms depend more on, and better use immigrant labor:

$$\Delta_b \equiv b_1 - b_0 > 0$$

This assumption stands in for a variety of reasons why firms that are more productive are more adept at using immigrants in production. They may be better at allocating immigrants and natives to different tasks, have better access to search networks where there are immigrant job-seekers, or may be less discriminatory toward (suffer less distaste from hiring) immigrant workers.

The cost function is given by

$$\left(\frac{c_j}{z}\right) Q_j(z) + c_j \kappa_j^f$$

where

$$c_j \equiv \left[a_L^{\sigma_E} \left(c_{Lj}\right)^{1-\sigma_E} + (w_S)^{1-\sigma_e}\right]^{\frac{1}{1-\sigma_E}}$$

and

$$c_{Lj} \equiv \left[b_j^{\sigma_I} \left(w_{IL}\right)^{1-\sigma_I} + (w_{NL})^{1-\sigma_I}\right]^{\frac{1}{1-\sigma_I}}$$

and w_{ie} represents the wage for a worker nativity i and education e and $w_S \equiv w_{NS} = w_{IS}$ because of perfect substitutability among higher-educated workers. κ_j^f is a fixed operating cost that is allowed to vary by technology choice for reasons mentioned in the text—immigrant-linked firms often appear to pay additional fixed operational costs in order to utilize immigrant labor. Thus, in order to access the immigrant-specific production boost represented by $b_1 > b_0$, they must pay a proportional cost every period, τ, such that $\kappa_1^f = \tau\kappa_0^f$.

The cost function leads to a familiar pricing rule in models of monopolistic competition and CES preferences:

$$p_j(z) = \left(\frac{\mu}{\mu - 1}\right)\left(\frac{c_j}{z}\right) \tag{C.2}$$

That is, the firm still charges a constant markup over its marginal cost, but the firm's marginal cost reflects the two different technology it uses. Firms compete through prices, and so firms that are able to pass on declines in c_j to consumers through p are able to gain in market share.

C.2 Equilibrium Conditions

Entrepreneurs only stay in the market if they are profitable. This defines a cutoff productivity for $j = 0$ firms:

$$\pi_0(z_0^*) \equiv 0 \tag{C.3}$$

[62]There is more robust evidence for imperfect substitutability among low-education workers. See, e.g., Peri and Sparber (2009).

A second cutoff exists, at which marginal producers are indifferent between the immigrant-intensive ($j = 1$) and non-immigrant-intensive ($j = 0$) production mode:

$$\pi_0(z_1^*) \equiv \pi_1(z_1^*) \tag{C.4}$$

Entrepreneurs with productivities below z_0^* exit the market, entrepreneurs with productivities in $[z_0^*, z_1^*]$ produce with technology 0, and entrepreneurs with productivities above z_1^* produce with technology 1. Entrepreneurs do not know their z prior to entry, and must pay an entry cost. The next equilibrium condition is free entry:

$$\mathbb{E}[\pi(z)] = \mathbb{E}[\pi(z)|z > z_0^*]\mathbb{P}[z > z_0^*] = c_0\kappa^e \tag{C.5}$$

where κ^e is a sunk (entry) cost entrepreneurs pay to take productivity draws, denominated in units of output. When profits are high enough, entrepreneurs enter until they no longer expect to recover their entry costs.[63] The price level, P is given by

$$P \equiv n_e \left[\int_{z_0^*}^{z_1^*} p_0(z)^{1-\mu} g(z)dz + \int_{z_1^*}^{\infty} p_1(z)^{1-\mu} g(z)dz\right] \tag{C.6}$$

where n_e is the endogenous mass of entrepreneurs who take productivity draws. Consumer spending Y is set equal to labor payments, and the final equilibrium conditions occur in the labor market, setting labor supply and labor demand equal for low-education immigrant and native workers. High-education worker wages are set to be the numeraire, $w_S \equiv 1$.

C.3 Equilibrium

The key items of interest revolve around z_0^*. First, I define

$$R_z \equiv \frac{z_1^*}{z_0^*} = \left[(c_0)^\mu \left(\frac{c_1\tau - c_0}{(c_1)^{1-\mu} - (c_0)^{1-\mu}}\right)\right]^{\frac{1}{\mu-1}} \tag{C.7}$$

and

$$\theta \equiv 1 + R_z^{-\phi}\left(\frac{c_1\tau - c_0}{c_0}\right) \tag{C.8}$$

Solving Equations (C.3) through (C.6) then yield

$$z_0^* = m\left[\left(\frac{\kappa_0^f}{\kappa^e}\right)\left(\frac{\mu - 1}{\phi - (\mu - 1)}\right)\theta\right]^{\frac{1}{\phi}} \tag{C.9}$$

$$F = Y\left(\frac{1}{c_0\kappa_0^f}\right)\left(\frac{\phi - (\mu - 1)}{\phi\mu}\right) \tag{C.10}$$

The key variable in this solution is θ, which sets Equation (C.9) apart from standard productivity cutoff expressions derived in similar models (e.g., Melitz, 2003). It introduces the notion that entry and exit decisions for marginal $j = 0$ firms depend on inframarginal $j = 1$ firms, through their

[63] Note that the assumption that entry costs scale with c_0 (instead of c_1 or a combination) is mostly made for analytical convenience. However, a simple, plausible justification is that producers do not invest in the costs to access immigrant labor until after entry activities have been completed and they find out they have a draw of z above z_1^*. Thus, the entry activities are paid for using type 0 technology.

ability to steal away market share when their costs go down. If c_1 goes down by more than c_0 in response to a shock, θ rises, which causes z_0^* to rise as well. The rise in z_0^* forces marginal type 0 firms to exit the market.

This mechanism drives the results below because of how it relates to the labor market. Section C.5 derives equilibrium in the labor market, with the upshot that demand curves slope downward. Thus, when a low-education tilted inflow of immigrants occurs, the wages of less-educated immigrants deteriorate the most of any group. In turn, c_1 falls by more than c_0 because $b_1 > b_0$.

C.3.1 *Value Added of the Model: P*

Price index P is inversely proportional to welfare. With z_0^* in hand, we can show[64]

$$P^{1-\mu} = [\text{Const.}](c_0)^{1-\mu} F^{\eta} (z_0^*)^{\mu-1} \theta$$

where F stands for firm mass. We then have

$$-\frac{d\log(P)}{dI} = \underbrace{-\frac{d\log(c_0)}{dI}}_{\text{Analogous to rep. firm model}} + \underbrace{\left(\frac{\eta}{\mu-1}\right)\frac{d\log(F)}{dI}}_{\text{Increased variety through more firms}} + \underbrace{\frac{d\log(z_0^*)}{dI}}_{\text{Culling of marginal firms}} + \underbrace{\left(\frac{1}{\mu-1}\right)\frac{d\log(\theta)}{dI}}_{\text{Technology switching}} \tag{C.11}$$

This expression clarifies the value added of this modeling framework. First, in a canonical, representative firm model of production in which all firms have access to $j = 0$ technology, the welfare impact of immigrants on native workers through prices would be defined only by the first term of Equation (C.11). This expression contains three additional avenues that this canonical framework misses. Note that I use the relationship between θ and z_o^* to combine the third and fourth terms in the main text.

The next two terms are driven by extensive margin changes: firm mass and the productivity level at the shut down margin. Sections 3 and 4 each delivered evidence that these reduced form parameters are positive in sign in partial equilibrium. Setting $\Delta_b = 0$, τ =1, and $c_0 = c_1$ implies a model that simply meld monopolistic competition in the output market with imperfect substitutability in the labor market. In such a model, Equations (C.7) through (C.9) show that $\theta = 1$ and that z_0^* is therefore constant. Thus, such a model would open up an avenue from firm mass to welfare through product variety—as long as consumers demand variety ($\eta = 1$)—but would not affect the firm productivity distribution. A model with $\Delta_b > 0$ and $\tau > 1$ opens the door to additional increases in welfare through a rising θ—which increases welfare both independently and through z_0^*. Simulation results below separate these channels.

C.4 Simulations

C.4.1 *Calibration*

Table A11 shows key model calibrations, set to match the U.S. economy in 2000, which had an immigrant share of 0.12 (55 percent of which has at most a high school degree) and 0.07 establishments per employee. Two calibrations are particularly difficult without employer-employee linked data. The first is the difference between b_1 and b_0, Δ_b. I will show results that vary this difference in order to demonstrate how big it has to be for θ to rise and for there to be productive realloca-

[64] $[\text{Const.}] = \left(\frac{\mu}{\mu-1}\right)^{1-\mu}\left(\frac{\phi}{\phi-(\mu-1)}\right)$

tion across firms. For a given Δ_b, b_0 is pinned down by the immigrant-native wage gap among low-education workers.

The second difficult parameter is τ, which controls the cost firms pay to obtain Δ_b, and determines how selected on productivity firms that pay it end up being. I also vary τ in simulations rather than set an arbitrary calibration, but this process revealed that $\tau < 1.5$ leads to unstable simulations and $\tau \geq 1.6$ leads to simulations that are essentially indistinguishable, when all other parameters are held fixed. I thus set τ to 2 for the figures presented below (see Table A12).

Under these calibrations, I run an experiment that increases the immigrant stock in the workforce in a specific way that matches the empirical content of the immigrant inflows in this paper. Each additional immigrant represents 0.69 additional high-school-equivalent immigrants and 0.31 college-equivalent immigrants (see Figure 2). As with a general immigrant inflow to the U.S. over the last 30 years, this experiment tilts the labor force towards lower educational attainment workers.

C.4.2 Results

Figure A8 shows the results from an example simulation with $\Delta_b = 0.3$ and $\tau = 2$. The top left figure shows that labor costs fall more for $j = 1$ firms. This mechanism filters through to the rest of the results: because $j = 1$ firms see lower marginal costs, the cutoff for switching to $j = 1$ technology moves down. The same is not true for $j = 0$ firms, even through immigrant entry does lower their production costs. This is because $j = 1$ firms are able to price compete away market share, leading to a higher productivity bar for $j = 0$ firms to be able to stay in the market. The result is more entry overall (of firms producing with both technology types), exit by marginal $j = 0$ firms, and an increase in native welfare. An immigration shock equivalent to a one percent increase in the population generates a 0.49 percent increase in native welfare. We can then use (C.11) to decompose this effect results into its component parts. Specifically,

$$\begin{aligned}
\mathcal{W}_I &\equiv \frac{d\log(\text{Real Native Income})}{dI} \\
&= \frac{d\log(w_{NU}N_U + w_S N_S)}{dI} - \frac{d\log(P)}{dI} \\
&= \underbrace{\frac{d\log(w_{NU}N_U + w_S N_S)}{dI} - \frac{d\log(c_0)}{dI}}_{\text{Standard}} + \underbrace{\left(\frac{\eta}{\mu-1}\right)\frac{d\log(F)}{dI}}_{\text{Variety}} + \underbrace{\frac{d\log(z_0^*)}{dI}}_{\text{Culling}} + \underbrace{\left(\frac{1}{\mu-1}\right)\frac{d\log(\theta)}{dI}}_{\text{Switching}}
\end{aligned}$$

where N_L and N_S are the fixed stock of native workers with at most a high school degree and more than a high school degree, respectively. This expression simply adds changes to native nominal income, $(w_{NL}N_L + w_S N_S)$, to Equation (C.11) .

The results of model simulations over a range of Δ_b are captured in Figure A9.[65] Three important findings emerge: first, the "standard" portions of the immigrants surplus that accrue to natives through wage and price changes, $\frac{d\log(w_{NL}N_L+w_S N_S)}{dI} - \frac{d\log(c_0)}{dI}$ (in blue), are a relatively small component of the immigrant surplus when we account for firm heterogeneity—less than 20 percent in the simulations conducted here. Gains from variety through increased firm presence, $\left(\frac{1}{\mu-1}\right)\frac{d\log(F)}{dI}$ (in gray), are the largest component of the immigrant surplus—insofar as consumers value variety. See Table A13 for more details on the decomposition.

[65] Note that when $\Delta_b = 0$, τ is set to 1. Otherwise, $\tau = 2$ as usual.

However, productive reallocation across firms, $\frac{d\log(z_0^*)}{dI} + \left(\frac{1}{\mu-1}\right)\frac{d\log(\theta)}{dI}$ (in red), particularly through the culling of marginal firms, $\frac{d\log(z_0^*)}{dI}$ (in darker red), plays a significant role as well—at least 30 percent of the immigrant surplus comes from productive reallocation. Returning to a primary motivation in this paper, these general equilibrium reallocation effects stem from the supply side of the product market—they occur due to immigrant characteristics as employees (making labor costs cheaper) and because they induce and increase entrepreneur mass n_e, which ultimately raises z_0^* through increased competition. Even if we are to believe that gains through variety are over-stated in models of monopolistic competition, taking firm heterogeneity into account is critical to understanding how immigration affects the economy. Specifically, the bifurcation in productivity that is generated by allowing firms to pay a cost to better-utilize immigrant workers more than doubles the native welfare gain that is associated with immigration.

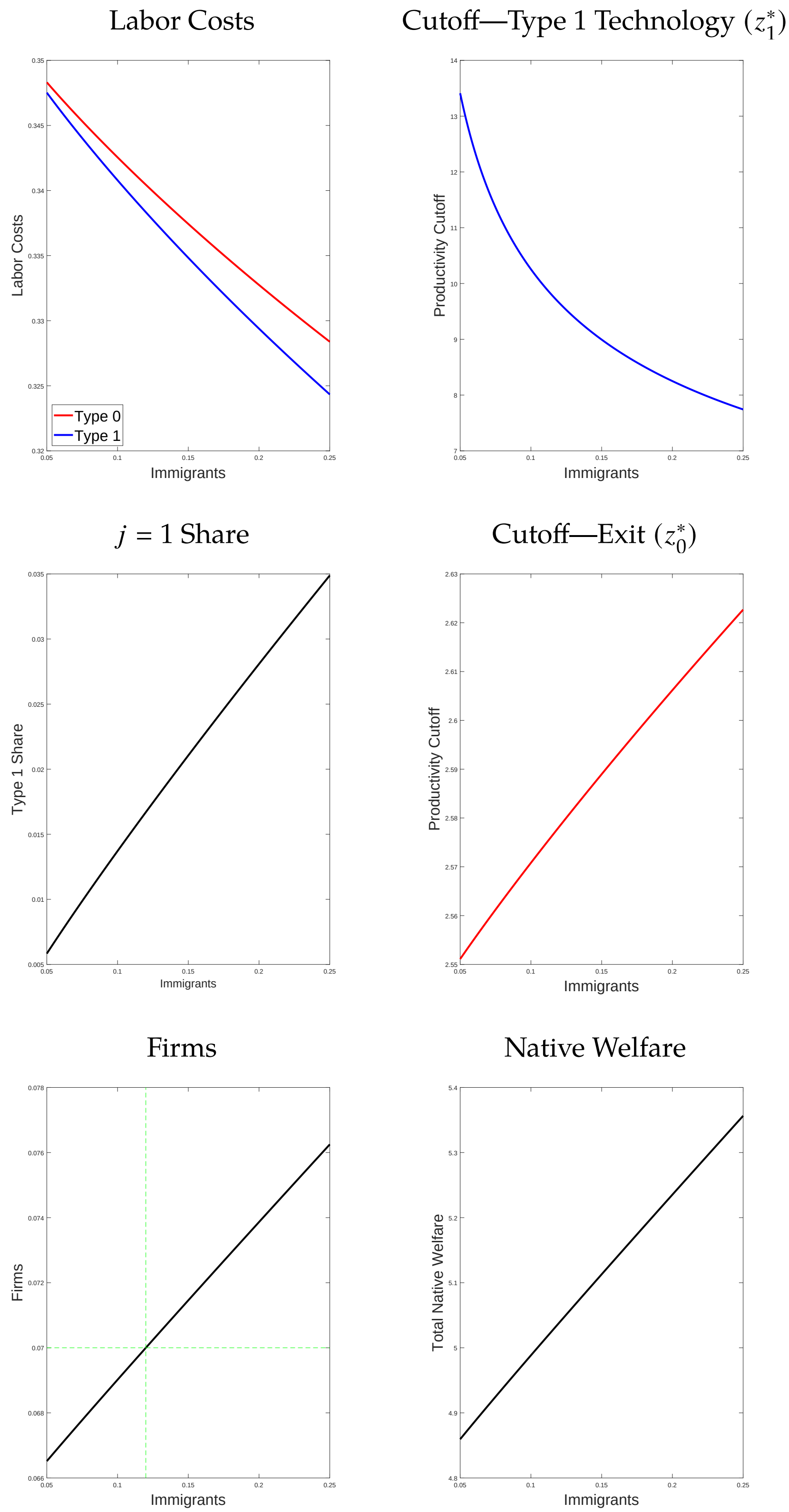
Table A11: Key Calibrations

Parameter	Value	Target Moments	Source
Panel A: Individually Calibrated			
σ_E	1.5	—	Ottaviano and Peri (2012)
σ_I	10	—	Ottaviano and Peri (2012)
μ	4	Average U.S. Markup = 32%	Christopoulou and Vermeulen (2012)
ϕ	3.1	$\phi > \mu - 1$	—
κ^e	3	—	—
m	1	—	—
Panel B: Jointly Calibrated			
a	0.64	$\frac{w_{NU}}{w_S} = 0.52$	2000 Census
b_0	$[0.15, 0.55]$	$\frac{w_{IU}}{w_S} = 0.4$	2000 Census
κ_0^f	0.25	$F = 0.05$	2000 Business Dynamics Statistics

Table A12: Percent Change in Native Welfare for a One Percent Increase in Workforce Due to Immigration

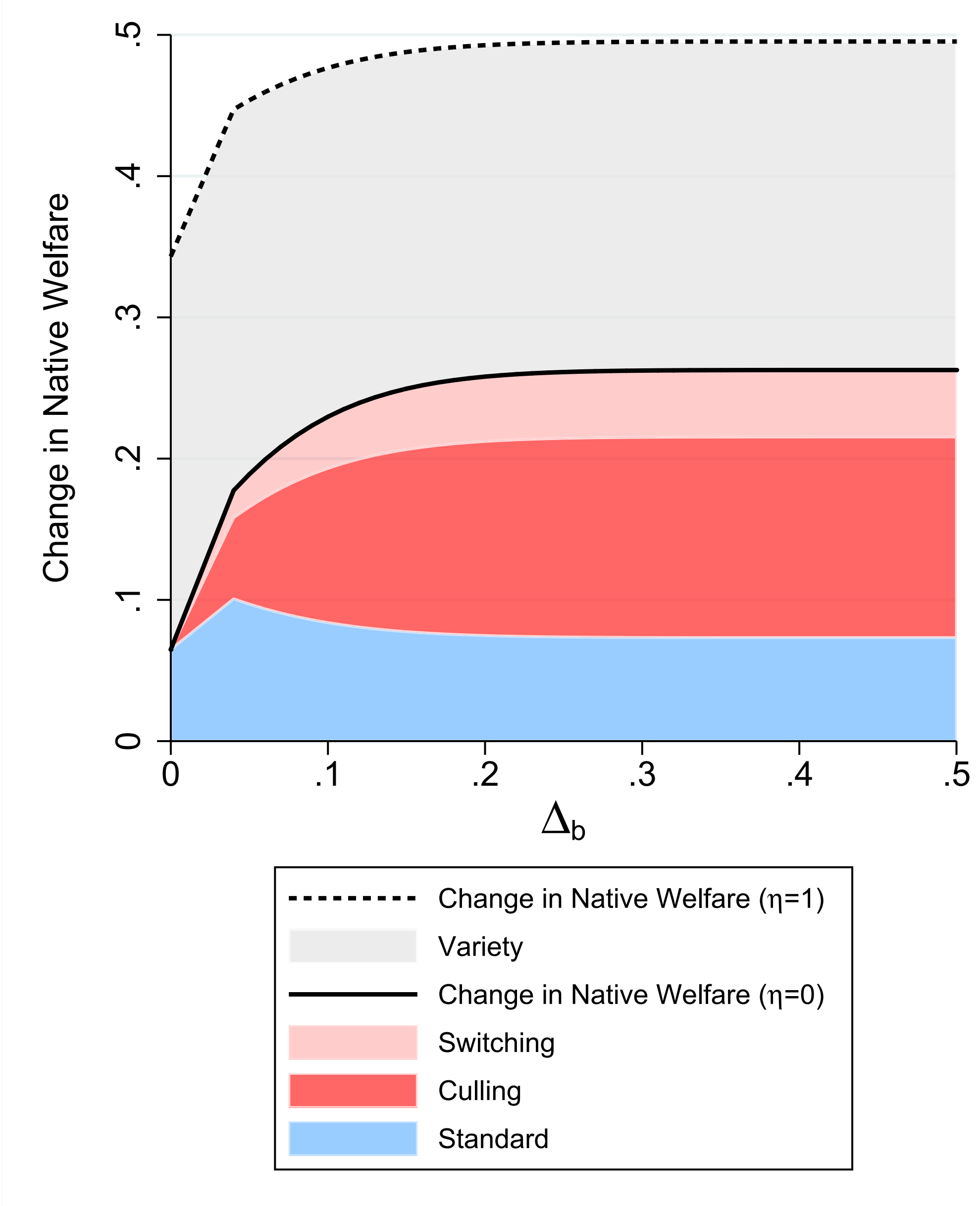
		Δ_b				
		0.1	0.2	0.3	0.4	0.5
	1.6	0.479	0.495	0.497	0.498	0.498
	1.7	0.479	0.495	0.497	0.498	0.498
τ	1.8	0.479	0.495	0.497	0.498	0.498
	1.9	0.479	0.495	0.498	0.498	0.498
	2	0.479	0.495	0.498	0.498	0.498

Figure A8: Example Simulation with $\Delta_b = 0.2$ and $\tau = 2$



Notes: Green dashed lines indicated data moments that were targeted in calibration.

Figure A9: Simulations—Percent Change in Native Welfare from a 1% Immigration Shock



Notes: When $\Delta_b = 0$, $\tau = 1$; otherwise, $\tau = 2$.

Table A13: Components of Immigrant Surplus

	Δ_b				
	0.1	0.2	0.3	0.4	0.5
Standard: $\frac{d\log(w_{NU}N_U+w_S N_S)}{dI} - \frac{d\log(c_0)}{dI}$	0.176	0.152	0.149	0.148	0.148
Firm Mass ($\eta = 1$): $\left(\frac{\eta}{\mu-1}\right)\frac{d\log(F)}{dI}$	0.518	0.476	0.470	0.470	0.470
Culling: $\frac{d\log(z_0^*)}{dI}$	0.230	0.279	0.286	0.286	0.286
Switching: $\left(\frac{1}{\mu-1}\right)\frac{d\log(\theta)}{dI}$	0.077	0.093	0.095	0.095	0.095
Total:	1	1	1	1	1

C.5 Derivation of Labor Market Equilibrium

In this section, I derive labor market equilibrium for low-education immigrant labor. An analogous derivation gives us labor market equilibrium for low-education native labor. High-education labor's price is set to be the numeraire, as described above. Firm's maximize the following expression for profits:

$$\pi_j(z) = p_j(z)q_j(z) - \sum_i \sum_e w_{ie} i_e - c_j \kappa_j^f$$

where $i \in \{I, N\}$ and $e \in \{U, S\}$. They take their demand curves, $q_j(z) = p_j(z)^{-\mu} P^{\mu-1} Y$ and their production functions $q_j(z) = zL_j(z)$ as given. Thus, first order conditions yield:

$$\begin{aligned} w_{IU} &= c_j L^{1/\sigma_E} a U_j^{1/\sigma_I - 1/\sigma_E} b_{Uj} I_{Uj}^{-1/\sigma_I} \\ w_{NU} &= c_j L^{1/\sigma_E} a U_j^{1/\sigma_I - 1/\sigma_E} N_{Uj}^{-1/\sigma_I} \end{aligned} \tag{C.12}$$

So, we have the familiar relative wage expression among low-education immigrant and native workers:

$$\frac{w_{IU}}{w_{NU}} = \left(\frac{I_{Uj}}{N_{Uj}}\right)^{-1/\sigma_I} b_{Uj}$$

Solving for N_{Uj} and plugging into the low-education aggregate U_j then yields the following expression for U_j in terms of I_{Uj}:

$$U_j = I_{Uj} b_{Uj}^{-\sigma_I} w_{IU}^{\sigma_I} c_{Uj}^{-\sigma_I}$$

Plugging this expression back into (C.12) yields the following expression for I_{Uj} in terms of wages and L_j:

$$I_{Uj} = w_{IU}^{-\sigma_I} b_{Uj}^{\sigma_I} c_j^{\sigma_E} c_{Uj}^{\sigma_I - \sigma_E} a^{\sigma_E} L_j \equiv I_{Uj}^{\text{unit}} L_j$$

Firms use L_j to produce output and to cover their fixed costs. Thus, integrating across firms yields the following expression that equates low-education immigrant labor supply, I_U, with labor demand:

$$I_U = n_e \left[\int_{z_0^*}^{z_1^*} I_{U0}^{\text{unit}} \left(\frac{q_0^*(z)}{z} + \kappa_0^f \right) g(z) dz + \int_{z_1^*}^{\infty} I_{U1}^{\text{unit}} \left(\frac{q_1^*(z)}{z} + \kappa_1^f \right) g(z) dz \right]$$

where $q_0^*(z)$ and $q_1^*(z)$ are optimal output choices—plugging in the pricing rule $p_j(z) = \left(\frac{\mu}{\mu-1}\right)\left(\frac{c_j}{z}\right)$ into the demand expression. After some algebra, the final expression for I_U becomes:

$$\begin{aligned} I_U = a^{\sigma_E} w_{IU}^{-\sigma_I} c_0^{-1} Y \Bigg\{ & b_{U0}^{\sigma_I} c_0^{\sigma_E} c_{U0}^{\sigma_I - \sigma_E} \left[\left(\frac{\mu}{\mu - 1}\right)^{-1} (\theta)^{-1} \left(1 - R_z^{-(\phi - (\mu - 1))}\right) + \left(\frac{\phi - (\mu - 1)}{\phi\mu}\right)\left(1 - R_z^{-\phi}\right) \right] + \\ & b_{U1}^{\sigma_I} c_1^{\sigma_E} c_{U1}^{\sigma_I - \sigma_E} \left(\frac{c_1}{c_0}\right)^{-\mu} \left[\left(\frac{\mu}{\mu - 1}\right)^{-1} (\theta)^{-1} \left(R_z^{-(\phi - (\mu - 1))}\right) + \tau \left(\frac{(\phi - (\mu - 1))}{\phi\mu}\right) \left(R_z^{-\phi}\right) \right] \Bigg\} \end{aligned}$$

D Additional Notes Related to Section 5

D.1 Comparing the Immigrant Surplus in a Representative Firm Model to a Model with Firm Heterogeneity

As a starting point, consider the following, constant elasticity of substitution (CES), model of production for a representative firm[66] in a local economy:

$$Q = \bar{z}\left(aI^{\frac{\sigma_I - 1}{\sigma_I}} + \bar{N}^{\frac{\sigma_I - 1}{\sigma_I}}\right)^{\frac{\sigma_I}{\sigma_I - 1}} \tag{D.1}$$

where I represents immigrant labor, $\bar{N}$ represents a fixed stock of native labor, $\bar{z}$ is total factor productivity, and σ_I is the elasticity of substitution between immigrant and native employees. While a more realistic version of the production function would start with nests for education and eventually work its way down to nativity (see, e.g., Ottaviano and Peri, 2008), this model delivers much of the intuition for the rest of this section in a simple way. In this context, σ_I can be thought of as a reduced form parameter that aggregates all the different reasons why an average immigrant worker may be different than an average native worker (including average educational attainment (see Figure 2).

Let native wages, w_N, be the numeraire. There is only one good, so consumer welfare is simply a function of its price and wages. When I increases in this model, the economy expands automatically (we can think of this as long run capital adjustment with a fixed rental rate in the background), but what matters to native workers is how it affects the price of the good. Denoting immigrant wages as w_I, P as the price of the good, taking first order conditions, and rearranging yields the following expression:

$$\mathcal{W}_I \equiv -\frac{d\log(P)}{dI} = -\frac{d\log(c)}{dI} \tag{D.2}$$

where

$$c \equiv (a^{\sigma_I} w_I^{1-\sigma_I} + 1)^{\frac{1}{1-\sigma_I}}$$

are labor costs to the firm. $\mathcal{W}_I$ is proportional to the immigrant surplus—the surplus accruing to the native workers as a result of immigrant inflows. Here, it is directly tied to the labor cost savings that occur as immigrant wages decline in response to I rising. Using similar, CES production function models, both Borjas (2014) and Ottaviano and Peri (2008) find that $\mathcal{W}_I$ is small and positive.

Sections 3 and 4 show that immigrants have large and heterogeneous effects on extensive margin decisions by firms. This suggests accounting for these effects in our theoretical analysis of immigration. This point has been made explicitly by di Giovanni et al. (2014), who feature gains from variety in their global welfare analysis of immigration. Melitz (2003) offers a simple way to incorporate extensive margin firm responses in long-run, steady state, general equilibrium analysis. The key features of this model are 1) consumer taste for variety, 2) a non-degenerate total factor productivity distribution across firms, and 3) a fixed cost of production. Combined, these features generate monopolistic competition that results in a non-trivial, but finite, firm mass, with heterogeneity across each firm's total factor productivity level.

We can gain simple insights into how our analysis of immigration may change when we account for these features by placing the production function from (D.1) into the closed economy Melitz

[66] Or, the aggregation of many, small, identical firms.

(2003) framework, where each firm is indexed by its own total factor productivity, z:

$$q(z) = zL(z)$$
$$L(z) = \left[a\,(I(z))^{\frac{\sigma_I - 1}{\sigma_I}} + (N(z))^{\frac{\sigma_I - 1}{\sigma_I}}\right]^{\frac{\sigma_I}{\sigma_I - 1}}$$

Consumers have elasticity of substitution μ across firms, leading to a conventional pricing rule for each firm:

$$p(z) = \left(\frac{\mu}{\mu - 1}\right)\left(\frac{c}{z}\right) \tag{D.3}$$

where w_I is the immigrant wage and we once again set native wages w_N to be the numeraire. The overall price index is given by

$$P^{1-\mu} = n_e \int_{z^*}^{\infty} p(z)^{1-\mu} g(z) dz$$

where n_e is the entrepreneur mass and z^* is the cutoff productivity, below which these entrepreneurs suffer losses and therefore exit in the long run. $g(z)$ is the distribution of productivity across firms in the local economy. I follow convention in assuming it is Pareto:

$$g(z) \equiv \phi m^{\phi} z^{-\phi - 1}$$

The mass of firms in the local economy is simply $F \equiv n_e \int_{z^*}^{\infty} g(z)dz = n_e m^{\phi} (z^*)^{-\phi}$.

With these simple ingredients in place, we can derive the following expression:

$$\mathcal{W}_I \equiv \underbrace{-\frac{d\log(c)}{dI}}_{\text{Same as above}} + \underbrace{\left(\frac{1}{\mu - 1}\right)\frac{d\log(F)}{dI}}_{\text{Increased variety through more firms}} + \underbrace{\frac{d\log(z^*)}{dI}}_{\text{Productivity pass-through to prices}} \tag{D.4}$$

In this setup, the immigration surplus has two additional terms compared to the representative firm model, in Equation (D.2). The first represents welfare gains in the form of increased consumer variety. The second represents welfare gains that arise from an increase in the productivity bar that entrepreneurs must cross in order to operate in the market. As seen in Equation (D.3), firms with higher z charge lower prices to consumers in order to compete away market share from their competitors. Thus, when z^* rises and lower productivity firms exit the market, consumers benefit through lower prices. The signs of each of these additional reduced form parameters are explored in detail in the empirical analyses of this paper. Section 3 finds that increased exposure to immigrant workers generates an increase in firm presence in local labor markets, while Section 5.5.1 finds that low productivity firms are culled from the market—in the language of this model, the latter indicates an increase in z^*.

D.2 Relationship to Equation (2.1) in a Model with Industry Groups

This section makes slight modifications to the model in Section (5) and above that help relate it to the partial equilibrium specifications employed in the empirical analyses.

Specifically, consider a commuting zone a closed economy and allow for K industries within each commuting zone. Consumers have Cobb-Douglas utility across sectors and CES preferences within sectors, which means that they spend a fixed share s_k on each given sector k. Labor endowments in each sector are initially fixed, and we are interested in the partial equilibrium comparative static that holds consumer demand fixed

$$\left.\frac{d\log(F_k)}{dI_k}\right|_{dY=0}$$

Fixing labor endowments is a simplifying assumption, but one that mimics the instrumental variable, which pushes immigrant workers into specific sectors based on comparative advantage. I also hold the native workforce in sector k fixed here. This precludes the empirical reality of native industry-switching in response to immigration. Thus, the exercise here can be thought of as deriving a version of Equation (2.1) with ΔN_{gkt} as a control (e.g., Column 4 of Table 2).

Under this setup and following the same steps as in Section (5) and above, we have

$$F_k = (s_k Y)\left(\frac{1}{c_{0k}\kappa_{0k}^f}\right)\left(\frac{\phi_k - (\mu_k - 1)}{\phi_k \mu_k}\right)$$
$$\Rightarrow \left.\frac{d\log(F_k)}{dI_k}\right|_{dY=0} = -\frac{d\log(c_{0k})}{dI_k}$$

a supply-side effect that generates extensive margin responses. However, a simple regression over of $\log(F_k)$ on I_k here would include the full effect:

$$\frac{d\log(F_k)}{dI_k} = -\frac{d\log(c_{0k})}{dI_k} + \frac{d\log(Y)}{dI_k}$$

However, $\frac{d\log(Y)}{dI_k}$ can be absorbed by a commuting zone fixed effect because it represents consumer spending across all sectors. For example, under our strict assumption of no mobility, $\frac{d\log(F_{k'})}{dI_k} = \frac{d\log(Y)}{dI_k}$. So,

$$\frac{d\log(F_k)}{dI_k} - \frac{d\log(F_{k'})}{dI_k} = -\frac{d\log(c_{0k})}{dI_k}$$

isolating the supply-side effect. This motivates the use of log changes (or DHS growth rates that well-approximate log changes but also allow for decompositions) and the inclusion of commuting-zone-by-year fixed effects in Equation (2.1).